TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG

Skripten der Mathematischen Institute

Wolfgang Marten

Lineare Algebra

für Studierende der
Informatik und Wirtschaftsinformatik

Wintersemester 2014/2015

Neunte, überarbeitete Auflage

Institut *Computational Mathematics*
AG Partielle Differentialgleichungen

Dr. Wolfgang Marten
Technische Universität Braunschweig
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen
Am Fallersleber Tore 1
D-38100 Braunschweig
Tel.: 0531 391 7404
E-Mail: `w.marten@tu-bs.de`

Lineare Algebra
für Studierende der
Informatik und Wirtschaftsinformatik

Wolfgang Marten
Technische Universität Braunschweig

Oktober 2014
Neunte, überarbeitete Auflage

Inhaltsverzeichnis

1 Einführendes Beispiel 5

2 Vektorräume. Basen. Lineare Abildungen 14

3 Inneres Produkt. Euklidische Vektorräume 31

4 Ergänzung: Hesse-Normalform 46

5 Matrizen. Die allgemeine lineare Gruppe 56

6 Kern und Bild einer Matrix. Gauß-Jordan 75

7 Koordinaten und Matrixdarstellungen 97

8 Diagonalisierung 102

9 Bestandteile der strikten Bruhat-Zerlegung 107

10 Strikte Bruhat-Zerlegung 121

11 LU-Zerlegung 132

12 Determinante. Bruhat. Laplace. Leibniz 134

13 Adjunkte und Regel von Cramer 149

14 Cayley-Hamilton. Leverrier-Faddeev 154

15 Matrixmodell der komplexen Zahlen 163

16 Spektralsatz. Projektionen. Gram-Schmidt 170

17 Ergänzung: Kleinste Fehlerquadrate. Gauß-Legendre **181**

18 Ergänzung: Cholesky-Zerlegung **186**

19 Ergänzung: Orthogonale Spiegelungen **189**

20 Drehungen. Euler-Winkel **191**

21 Ergänzung: Gruppen **210**

22 Ergänzung: Euklidische Bewegungen **215**

23 Ergänzung: Abbildungen **219**

24 Anhang: Klausuren **222**

Literatur **261**

Ab dem Wintersemester 2007/08 müssen die Studierenden der Wirtschaftsinformatik lediglich die erste Hälfte der Vorlesung hören. Diese erste Hälfte umfasst die Abschnitte 1 bis 8, 11, 23 und Teile von 12, 13, 14, 17.

In Anbetracht der üblichen Ausfallempfehlungen nimmt die erste Hälfte der Vorlesung sieben bis acht Wochen des Wintersemesters in Anspruch.

Von den ersten acht Auflagen des Skriptes wurden Druckfassungen an die Studierenden verteilt. Der Druck und die Bindung wurden aus Studiengebühren und aus dem Institutsetat finanziert. Die gedruckten Versionen enthielten einen kompletten Satz Übungsblätter aus einem früheren Wintersemester und alte Klausuren mit den Korrekturvorlagen.

Das Skript ersetzt weder die Teilnahme an den vorgesehenen Lehrveranstaltungen noch eine eingehende Nachbereitung. Die Beschränkung auf sogenannte klausurrelevante Themen führt in der Regel zu mangelhaften Studienerfolgen. Zur Nachbereitung gehören beispielsweise das Nachvollziehen der Beweise und das Nachrechnen der Beispiele.

1 Einführendes Beispiel

Wir beginnen mit der Erörterung eines einfachen linearen Gleichungssystems, das von einer geometrischen Fragestellung ausgeht. Bei dieser Gelegenheit werden wir die Verfahren von *Gauß* und *Gauss-Jordan* in Aktion sehen und einen Spezialfall der Regel von *Cramer* kennenlernen.

Eine Gerade ist durch die Angabe eines beliebigen Paares zweier verschiedener Punkte, die auf der Geraden liegen, eindeutig bestimmt. Seien

$$\begin{aligned} g_1 : &\qquad \text{Gerade durch } (3,0) \text{ und } (-4,6)\,, \\ g_2 : &\qquad \text{Gerade durch } (1,-1) \text{ und } (6,-4) \end{aligned}$$

gegeben. Einer Skizze entnehmen wir, dass beide Geraden nicht zusammenfallen und auch nicht parallel sind. Also gibt es genau einen Schnittpunkt. Gesucht sind die Koordinaten x_1 und x_2 des Schnittpunktes.

Wir beschreiben die beiden Geraden g_1 und g_2 durch lineare Gleichungen:

$$\begin{aligned} g_1 : &\qquad 6x_1 + 7x_2 = 18\,, \\ g_2 : &\qquad 3x_1 + 5x_2 = -2\,. \end{aligned}$$

Die Koordinaten x_1 und x_2 des Schnittpunktes von g_1 und g_2 sind Lösungen des *inhomogenen linearen Gleichungssystems*

$$\begin{aligned} 6x_1 + 7x_2 &= 18\,, \\ 3x_1 + 5x_2 &= -2\,. \end{aligned}$$

Wir lösen das Gleichungssystems durch *elementare Zeilenumformungen* und *Rückwärtselimination*. Nach den geometrischen Vorbetrachtungen vermuten wir, dass es genau eine Lösung (x_1, x_2) gibt.

$$\begin{aligned} 6x_1 + 7x_2 &= 18 \\ 3x_1 + 5x_2 &= -2 \\ -\,-\,-\,-\,-\,-&-\,-\,-\,-\,-\,- \\ 6x_1 + 7x_2 &= 18 \\ \tfrac{3}{2}x_2 &= -11 \qquad\qquad Z_2 - \tfrac{1}{2}Z_1 \end{aligned}$$

Rechts neben der Rechnung haben wir den Typ der elementaren Zeilenumformung als Kommentar notiert. Dabei beziehen sich Z_1 und Z_2 auf die erste respektive zweite Zeile des vorherigen Schrittes oberhalb der gestrichelten Linie. Die Hälfte der ersten Zeile Z_1 ist von der zweiten Zeile Z_2 abgezogen worden. Die erste Zeile Z_1 ist unverändert übernommen worden.

Durch eine elementare Zeilenumformung ist das Gleichungssystem in eine *Stufenform* gebracht worden, aus der sich zuerst x_2 und dann x_1 berechnen lässt:

$$\begin{aligned} \tfrac{3}{2}x_2 &= -11 \\ \Longrightarrow \quad x_2 &= -\tfrac{22}{3} \end{aligned}$$

$$\Longrightarrow \quad 6x_1 + 7 \cdot (-\tfrac{22}{3}) = 18$$

$$\Longrightarrow \quad 6x_1 = \tfrac{54+154}{3} = \tfrac{208}{3}$$

$$\Longrightarrow \quad x_1 = \tfrac{104}{9}\,.$$

Diese sukzessive Berechnung von x_2 und x_1 heißt *Rückwärtselimination.*

Wir überlegen uns, was wir bisher erreicht haben. Unter der Annahme, dass das (x_1, x_2) eine Lösung ist, haben wir gezeigt, dass

$$(x_1, x_2) = (\tfrac{104}{9}, -\tfrac{22}{3})$$

gelten muss. Eine Probe zeigt, dass wir tatsächlich eine Lösung erhalten haben.

$$6 \cdot \tfrac{104}{9} + 7 \cdot (-\tfrac{22}{3}) = \tfrac{208-154}{3} = \tfrac{54}{3} = 18\,,$$

$$3 \cdot \tfrac{104}{9} + 5 \cdot (-\tfrac{22}{3}) = \tfrac{104-110}{3} = -\tfrac{6}{3} = -2\,.$$

Damit haben wir abschließend gezeigt, dass

$$(x_1, x_2) = (\tfrac{104}{9}, -\tfrac{22}{3})$$

die einzige Lösung des Gleichungssystems ist.

Die Rechnung lässt sich in zweierlei Hinsicht vereinfachen. Wir notieren lediglich das Koeffizientenschema und ersetzen die Rückelimination durch weitere elementare Zeilenumformungen bis wir die *Treppennormalform* des Koeffizientenschemas erreicht haben. Weil alle vorgenommenen Zeilenumformungen rückgängig gemacht werden können, erhalten wir ein lineares Gleichungssystem, das dieselben Lösungen wie das ursprünglich gegebene besitzt.

Hier sei schon darauf hingewiesen, dass uns die Treppennormalform später mehrfach begegnen wird. Die Treppennormalform eines Koeffizientenschemas oder einer *Matrix*, wie wir sagen werden, hängt nicht von den vorgenommenen Zeilenumformungen ab.

Wir führen den angekündigten vereinfachten Rechengang zur Lösung des inhomogenen linearen Gleichungssystems im Einzelnen vor.

$$\begin{array}{cc|c||c}
6 & 7 & 18 & \\
3 & 5 & -2 & \\
\hline
6 & 7 & 18 & \\
0 & \frac{3}{2} & -11 & Z_2 - \frac{1}{2}Z_1 \\
\hline
1 & \frac{7}{6} & 3 & \frac{1}{6}Z_1 \\
0 & 1 & -\frac{22}{3} & \frac{2}{3}Z_2 \\
\hline
1 & 0 & \frac{104}{9} & Z_1 - \frac{7}{6}Z_2 \\
0 & 1 & -\frac{22}{3} &
\end{array}$$

Rechts neben dem vertikalen Doppelstrich notieren wir die vorgenommenen elementaren Zeilenumformungen. Die letzte Spalte der Treppennormalform enthält die Lösung des Gleichungssystem. Die Rückwärtselimination reduziert sich in dieser Situation auf ein bloßes Ablesen.

Das zusammenfassende Schema des Rechenganges legt nahe, die Koordinaten des gesuchten Schnittpunktes als Komponenten in einen *Spaltenvektor* einzutragen. Entsprechend werden die Koeffizienten der rechten Seite des Gleichungssystems als Komponenten in einen Spaltenvektor und die Koeffizienten der linken Seite als Einträge in eine (2×2)-*Matrix* eingetragen. Das Gleichungssystem nimmt dann die Form

$$\begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 18 \\ -2 \end{pmatrix}$$

an. Die bloße Übersetzung eines Systems linearer Gleichungen in eine Matrixgleichung ist noch keine großartige Leistung. Entscheidend ist, dass mit Matrizen und Vektoren gerechnet werden kann.

Die Rechengesetze für Matrizen und Vektoren bilden eines der Themen der Linearen Algebra. Die Rechengesetze sind so gemacht, dass sie das Rechnen mit linearen Gleichungen wiedergeben und überschaubarer gestalten.

Wir setzen die Erörterung des Beispiels

$$\begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 18 \\ -2 \end{pmatrix}$$

fort. Auf der linken Seite der Gleichung steht ein Produkt einer (2×2)-Matrix mit einem 2-komponentigen Spaltenvektor. Das Ergebnis auf der rechten Seite ist ebenfalls ein 2-komponentiger Spaltenvektor. Wir formulieren die Rechenregel für diesen Spezialfall. Die Komponenten b_i des Spaltenvektors

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1.1}$$

sind durch die Gleichungen

$$b_i = \sum_{j=1}^{2} a_{ij} x_j \, , \quad i = 1, 2 \tag{1.2}$$

definiert. Wir haben gesehen, dass die Gleichung

$$\begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 18 \\ -2 \end{pmatrix}$$

genau eine Lösung besitzt. Kann diese Lösung durch Inversion der Matrix

$$\begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix}$$

berechnet werden? Das heißt, gibt es eine (2×2)-Matrix mit Einträgen c_{ij} derart, dass

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} 18 \\ -2 \end{pmatrix}$$

gilt?

Satz 1.1. *Seien $a_{ij} \in \mathbb{R}$ und $b_i \in \mathbb{R}$ mit $i, j = 1, 2$ und*

$$d = a_{11}a_{22} - a_{21}a_{12} \neq 0 \tag{1.3}$$

gegeben. Dann gibt es eindeutig bestimmte $x_i \in \mathbb{R}$ mit $i = 1, 2$ und

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} . \tag{1.4}$$

Es gilt

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \dfrac{a_{22}}{d} & -\dfrac{a_{12}}{d} \\ -\dfrac{a_{21}}{d} & \dfrac{a_{11}}{d} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} . \tag{1.5}$$

Ausgeschrieben bedeutet dies

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}} , \tag{1.6}$$

$$x_2 = \frac{-a_{21}b_1 + a_{11}b_2}{a_{11}a_{22} - a_{21}a_{12}} . \tag{1.7}$$

Bemerkungen. Die Formeln (1.6) und (1.7) sind ein Spezialfall der *Regel von Cramer*. Wir werden sehen, dass

$$d = a_{11}a_{22} - a_{21}a_{12}$$

die *Determinante* der Matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

ist. Wir schreiben

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12} . \tag{1.8}$$

Die Bedingung $d \neq 0$ bedeutet, dass diese Matrix *invertierbar* ist. Die *inverse Matrix* ist

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{a_{22}}{d} & -\dfrac{a_{12}}{d} \\ -\dfrac{a_{21}}{d} & \dfrac{a_{11}}{d} \end{pmatrix} .$$

Bevor wir den Beweis des Satzes 1.1 behandeln, wenden wir den Satz auf das obige Beispiel an.

Beispiel 1.2.

$$\begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 18 \\ -2 \end{pmatrix}.$$

Der Satz ist anwendbar, denn es gilt

$$d = 6 \cdot 5 - 3 \cdot 7 = 30 - 21 = 9 \neq 0\,.$$

Damit erhalten wir unser früheres Ergebnis

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{9} & -\frac{7}{9} \\ -\frac{3}{9} & \frac{6}{9} \end{pmatrix} \begin{pmatrix} 18 \\ -2 \end{pmatrix} = \begin{pmatrix} \frac{90+14}{9} \\ \frac{-54-12}{9} \end{pmatrix} = \begin{pmatrix} \frac{104}{9} \\ -\frac{22}{3} \end{pmatrix}.$$

□

Beweis. Wir zerlegen den Beweis des Satzes 1.1 in drei Teilschritte. Jeder Teilschritt enthält typische Vorgehensweisen.

Erster Schritt. Zuerst rechnen wir nach, dass die Formeln

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}}\,, \quad x_2 = \frac{-a_{21}b_1 + a_{11}b_2}{a_{11}a_{22} - a_{21}a_{12}}$$

Lösungen des Gleichungssystems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1\,, \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

liefern. Einsetzen ergibt

$$a_{11}x_1 + a_{12}x_2 = \frac{a_{11}(a_{22}b_1 - a_{12}b_2) + a_{12}(-a_{21}b_1 + a_{11}b_2)}{a_{11}a_{22} - a_{21}a_{12}} = b_1\,,$$

$$a_{21}x_1 + a_{22}x_2 = \frac{a_{21}(a_{22}b_1 - a_{12}b_2) + a_{22}(-a_{21}b_1 + a_{11}b_2)}{a_{11}a_{22} - a_{21}a_{12}} = b_2\,.$$

Damit ist bewiesen, dass die Formeln tatsächlich Lösungen liefern. Der Nachweis der Einzigkeitsaussage benötigt eine Vorbereitung, die wir im zweiten Schritt treffen.

Zweiter Schritt. Wir überlegen uns, dass das zugehörige homogene lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0\,, \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

nur die Lösungen $x_1 = x_2 = 0$ besitzt. Nach Voraussetzung gilt

$$a_{11}a_{22} - a_{21}a_{12} \neq 0\,.$$

Also können a_{11} und a_{12} nicht zugleich verschwinden. Wir betrachten zuerst den Fall $a_{11} \neq 0$. Auflösen der ersten Gleichung nach x_1 ergibt

$$x_1 = -\frac{a_{12}}{a_{11}}x_2\,.$$

Einsetzen in die zweite Gleichung liefert

$$0 = a_{21}x_1 + a_{22}x_2 = -a_{21}\frac{a_{12}}{a_{11}}x_2 + a_{22}x_2 = \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}}x_2$$

Wegen $a_{11}a_{22} - a_{21}a_{12} \neq 0$ folgt

$$x_2 = 0\,, \quad x_1 = -\frac{a_{12}}{a_{11}}x_2 = 0\,.$$

Den Fall $a_{12} \neq 0$ behandeln wir analog. Auflösen der ersten Gleichung nach x_2 ergibt

$$x_2 = -\frac{a_{11}}{a_{12}}x_1\,.$$

Einsetzen in die zweite Gleichung ergibt

$$0 = a_{21}x_1 + a_{22}x_2 = a_{21}x_1 - a_{22}\frac{a_{11}}{a_{12}}x_1 = \frac{a_{21}a_{12} - a_{22}a_{11}}{a_{12}}x_1\,.$$

Wegen $a_{11}a_{22} - a_{21}a_{12} \neq 0$ folgt

$$x_1 = 0\,, \quad x_2 = -\frac{a_{11}}{a_{12}}x_1 = 0\,.$$

Damit ist die Behauptung über die Lösungen des zugehörigen homogenen linearen Gleichungssystems bewiesen.

Dritter Schritt. Nun beweisen wir die Einzigkeitsaussage des Satzes. Es seien $x_1, x_2, y_1, y_2 \in \mathbb{R}$ mit

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1\,, \\ a_{21}x_1 + a_{22}x_2 &= b_2\,, \\ a_{11}y_1 + a_{12}y_2 &= b_1\,, \\ a_{21}y_1 + a_{22}y_2 &= b_2 \end{aligned}$$

gegeben. Differenzbildung der ersten und dritten respektive der zweiten und der vierten Gleichung ergibt

$$\begin{aligned} a_{11}(x_1 - y_1) + a_{12}(x_2 - y_2) &= b_1 - b_1 = 0\,, \\ a_{21}(x_1 - y_1) + a_{22}(x_2 - y_2) &= b_2 - b_2 = 0\,. \end{aligned}$$

Nach den obigen Ausführungen folgt

$$x_1 = y_1\,, \quad x_2 = y_2\,.$$

Damit ist der Beweis des Satzes 1.1 beendet. □

Beispiel 1.3.

$$\begin{pmatrix} \frac{1}{5} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{7}{6} \\ 1 \end{pmatrix}\,.$$

$$
\begin{array}{cc|c||r}
\frac{1}{5} & \frac{1}{6} & \frac{7}{6} & \\
\frac{1}{6} & \frac{1}{7} & 1 & \\
\hline
1 & \frac{5}{6} & \frac{35}{6} & 5\,Z_1 \\
\frac{1}{6} & \frac{1}{7} & 1 & \\
\hline
1 & \frac{5}{6} & \frac{35}{6} & \\
0 & \frac{1}{252} & \frac{1}{36} & Z_2 - \frac{1}{6}Z_1 \\
\hline
1 & \frac{5}{6} & \frac{35}{6} & \\
0 & 1 & 7 & 252\,Z_2 \\
\hline
1 & 0 & 0 & Z_1 - \frac{5}{6}Z_2 \\
0 & 1 & 7 &
\end{array}
$$

$$
\frac{1}{5} \cdot \frac{1}{7} - \frac{1}{6} \cdot \frac{1}{6} = \frac{36-35}{1260} = \frac{1}{1260} \neq 0\,.
$$

$$
\begin{aligned}
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} \frac{1260}{7} & -\frac{1260}{6} \\ -\frac{1260}{6} & \frac{1260}{5} \end{pmatrix} \begin{pmatrix} \frac{7}{6} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 180 & -210 \\ -210 & 252 \end{pmatrix} \begin{pmatrix} \frac{7}{6} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 210-210 \\ -245+252 \end{pmatrix} = \begin{pmatrix} 0 \\ 7 \end{pmatrix}\,.
\end{aligned}
$$

Im Hinblick auf die Untersuchung von linearen Gleichungssystemen mit n Gleichungen in m Unbekannten verallgemeinern wir (1.1) und (1.2). Ein lineares Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1m}x_m &= b_m \,, \\ &\vdots \\ a_{n1}x_1 + \ldots + a_{nm}x_m &= b_m \end{aligned}$$

in den Unbekannten $x_1, \ldots, x_m$ schreiben wir in der Form

$$\sum_{j=1}^{m} a_{ij}x_j = b_i \,, \quad i = 1, \ldots, n \,. \tag{1.9}$$

Dafür schreiben wir

$$\begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} . \tag{1.10}$$

Die $(n \times m)$-Matrix (a_{ij}) mit

$$(a_{ij}) = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{nm} \end{pmatrix}$$

heißt die *Systemmatrix* von (1.9) respektive (1.10). Mit $\mathbb{R}^{n\times m}$ bezeichnen wir die Menge der reellen $(n \times m)$-Matrizen. Im Fall $n = m$ lässt sich für beliebiges $n \in \mathbb{N}$ die Regel von Cramer beweisen. Siehe 13.5. Dazu ist es erforderlich, die Determinante für $(n \times n)$-Matrizen zu definieren. Eine mögliche Definition ist die direkte Definition 1.4 durch Rekursion.

Definition 1.4. *Sei $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times n}$. Im Fall $n \geq 2$ sei $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ die quadratische Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Die Determinante* $\det(A)$ *der Matrix A wird rekursiv durch*

$$\det(A) = \begin{cases} a_{11} \,, \quad n = 1 \,, \\ a_{11}a_{22} - a_{21}a_{12} \,, \quad n = 2 \,, \\ \displaystyle\sum_{\nu=1}^{n} (-1)^{\nu+1} a_{\nu 1} \det(A_{\nu 1}) \,, \quad n \geq 2 \end{cases}$$

definiert.

In Abschnitt 12 werden wir die Determinante einer quadratischen Matrix, die invertierbar ist, mit Hilfe einer Bruhat-Zerlegung der Matrix definieren und auf diese Weise die allgemeinen Eigenschaften der Determinante herleiten. Die Invertierbarkeit und die strikte Bruhat-Zerlegung können simultan mit einer Abwandlung des Gauß-Jordan-Verfahrens bestimmt werden. Dieses Verfahren beruht auf elementaren Zeilenumformungen.

Beispiel 1.5.

$$\det \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix} = 2 \cdot 2 - 1 \cdot 0 + 2 \cdot 0 = 4\,.$$

Siehe Beispiel 8.9. □

2 Vektorräume. Basen. Lineare Abildungen

Die reellen Zahlen nennen wir auch *Skalare*. Produkte von Skalaren mit Spaltenvektoren und Summen von Spaltenvektoren werden komponentenweise definiert. Auf diese Weise werden bestimmte Rechenregeln von den reellen Zahlen auf die Spaltenvektoren übertragen. Damit erhalten wir den Raum $\mathbb{R}^n$ der Spaltenvektoren. Der Raum $\mathbb{R}_n$ der Zeilenvektoren wird analog definiert. In 2.10 geben wir Definitionen, die die spezielle Anordnungen der Komponenten als Spaltenvektoren respektive Zeilenvektoren vermeiden. Die vorläufige Definition des $\mathbb{R}^n$ als Spaltenraum dient zur Motivation des Vektorraumbegriffes in Definition 2.2.

Mit $\mathbb{R}^n$ bezeichnen wir die Menge der n-komponentigen Spaltenvektoren

$$x = (x_i) = (x_i)_{i=1,\dots,n} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

deren Komponenten $x_1, \dots, x_n$ reelle Zahlen sind. Für eine Definition, die diese spezielle Anordnung der Komponenten vermeidet, siehe 2.10. Die reelle Zahl x_i ist die *i-te Komponente* von x. In Formeln schreiben wir

$$\begin{aligned} \mathbb{R}^n &= \{ (x_i)_{i=1,\dots,n} \mid x_i \in \mathbb{R}\,,\ i = 1, \dots, n \} \\ &= \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R}\,,\ i = 1, \dots, n \}\,. \end{aligned}$$

Die Notation als Spaltenvektoren ist eine Konvention, die aus Gründen der Rechentechnik verabredet wird. Im laufenden Text ist oft günstiger, eine Zeilenschreibweise zu verwenden. Dies geschieht mit Hilfe der Transposition. Diese Operation überführt Zeilen in Spalten und Spalten in Zeilen.

$$(x_1, \dots, x_n)^t = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^t = (x_1, \dots, x_n)\,.$$

Doppelte Anwendung der Transposition ist die identische Operation.

$$(x_1, \dots, x_n)^{t\,t} = (x_1, \dots, x_n)\,, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^{t\,t} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Die Definition der Menge $\mathbb{R}^n$ können wir nun in der Form

$$\begin{aligned} \mathbb{R}^n &= \{ (x_i)_{i=1,\dots,n} \mid x_i \in \mathbb{R}\,,\ i = 1, \dots, n \} \\ &= \{ (x_1, \dots, x_n)^t \mid x_i \in \mathbb{R}\,,\ i = 1, \dots, n \} \end{aligned}$$

ausdrücken.

Nach diesen Vorbereitungen definieren wir die Multiplikation eines Skalars mit einem Spaltenvektor und die Addition zweier Spaltenvektoren. Für einen beliebigen Skalar $\gamma \in \mathbb{R}$ und beliebige Spaltenvektoren

$$x = (x_i)_{i=1,\ldots,n} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n\,, \quad y = (y_i)_{i=1,\ldots,n} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

definieren wir

$$\gamma x \equiv \gamma \cdot x \equiv (\gamma x_i)_{i=1,\ldots,n} = \begin{pmatrix} \gamma x_1 \\ \vdots \\ \gamma x_n \end{pmatrix} \in \mathbb{R}^n\,, \tag{2.1}$$

wobei wir im Allgemeinen das Multiplikationszeichen fortlassen, und

$$x + y \equiv (x_i + y_i)_{i=1,\ldots,n} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \in \mathbb{R}^n\,. \tag{2.2}$$

Nach Definition gilt

$$\gamma x_i = i\text{-te Komponente von } \gamma x\,,$$

$$x_i + y_i = i\text{-te Komponente von } x + y\,.$$

Trivialerweise gilt

$$1x = 1 \cdot x = x\,. \tag{2.3}$$

Wir heben ausdrücklich hervor, dass der Ausdruck $x\gamma$ nicht definiert worden ist. Ein Skalar wird nach Formel (2.1) von links an einen Spaltenvektor multipliziert. Durch Formel (2.2) wird lediglich die Addition zweier Spaltenvektoren mit derselben Anzahl von Komponenten definiert. Beispielsweise ist die Summe

$$\begin{pmatrix} -2 \\ 5 \end{pmatrix} + \begin{pmatrix} 78 \\ -1 \\ 23 \end{pmatrix}$$

durch (2.2) nicht definiert. In den Formeln (2.1) und (2.2) bezeichnen x_i und y_i die Komponenten von x respektive y.

In anderen Situationen bezeichnen wir mit x_i und y_i Spaltenvektoren oder andere mathematische Objekte. Was gemeint ist, geht aus der Angabe, in welcher Menge die x_i und y_i enthalten sind, hervor.

Beispiel 2.1. Gegeben seien der reelle Skalar $\gamma = 5$ und die Spaltenvektoren

$$x = \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix}, \quad y = \begin{pmatrix} 45 \\ 0 \\ -21 \\ 14 \end{pmatrix}$$

im $\mathbb{R}^4$. Dann gilt

$$\gamma x = 5 \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix} = \begin{pmatrix} -15 \\ 35 \\ 115 \\ -125 \end{pmatrix}, \quad x + y = \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix} + \begin{pmatrix} 45 \\ 0 \\ -21 \\ 14 \end{pmatrix} = \begin{pmatrix} 42 \\ 7 \\ 2 \\ -11 \end{pmatrix}.$$

□

Ein Spaltenvektor $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ mit $x_i = 0$ für $i = 1, \dots, n$ heißt *Nullvektor* und wird ebenfalls mit 0 bezeichnet.

$$0 \equiv (0, \dots, 0)^t = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n. \tag{2.4}$$

Sei $x \in \mathbb{R}^n$ ein beliebiger Spaltenvektor. Aus (2.1) und (2.4) folgt

$$0x = 0 \cdot x = 0. \tag{2.5}$$

Im linken und im mittleren Teil der Gleichungskette (2.5) wird der Spaltenvektor $x \in \mathbb{R}^n$ von links mit dem Skalar $0 \in \mathbb{R}$ multipliziert. Im rechten Teil steht der Nullvektor $0 \in \mathbb{R}^n$.

Mit dem Symbol 0 bezeichnen wir generell das *neutrale Element* bezüglich einer Addition. Nach (2.12) ist der Nullvektor das neutrale Element bezüglich der Addition von Spaltenvektoren. Neben dem Nullvektor sind die Nullfunktion, das Nullpolynom und die Nullmatrix weitere Beispiele.

Wir definieren den *negativen Vektor* $-x \in \mathbb{R}^n$ eines beliebigen Spaltenvektors $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ durch die Gleichung

$$-x \equiv (-1)x = \begin{pmatrix} (-1)x_1 \\ \vdots \\ (-1)x_n \end{pmatrix} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix} \in \mathbb{R}^n. \tag{2.6}$$

Dabei haben wir von der Rechenregel $-\alpha = (-1)\alpha$ für reelle Zahlen Gebrauch gemacht. Der Ausdruck $-x$ auf der linken Seite der Formel (2.6) ist als ein neues Symbol zu lesen. Zur Betonung dieser Lesart wird das Symbol $-x$ innerhalb einer Rechnung in Klammern eingeschlossen. Klammern werden gesetzt, um Reihungen der Form $--$ oder $+-$ zu vermeiden.

Im Sinne der Klammerkonvention gilt

$$-x + y = (-x) + y = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} -x_1 + y_1 \\ \vdots \\ -x_n + y_n \end{pmatrix} . \tag{2.7}$$

Aus (2.7) und (2.4) folgt als Spezialfall die Rechenregel

$$-x + x = 0 . \tag{2.8}$$

Die Summe von y und $-x$ schreiben wir im Sinne der Klammerkonvention in der Form

$$y + (-x)$$

nieder. Nach (2.6) gilt

$$y + (-x) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix} = \begin{pmatrix} y_1 + (-x_1) \\ \vdots \\ y_n + (-x_n) \end{pmatrix} = \begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{pmatrix} .$$

Die Definition $\alpha - \beta = \alpha + (-\beta)$ für die Differenzbildung reeller Zahlen übertragen wir auf das Rechnen mit Spaltenvektoren und definieren

$$y - x \equiv y + (-x) = \begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{pmatrix} . \tag{2.9}$$

Aus (2.9) und (2.4) folgt als Spezialfall die Rechenregel

$$x - x = 0 . \tag{2.10}$$

Aus der Formel $-(-\alpha) = \alpha$ für reelle Zahlen ergibt sich die analoge Formel für Spaltenvektoren.

$$-(-x) = -\begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix} = \begin{pmatrix} -(-x_1) \\ \vdots \\ -(-x_n) \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x . \tag{2.11}$$

Wir wenden uns noch einmal dem Nullvektor zu. Aus (2.2) folgt

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 + 0 \\ \vdots \\ x_n + 0 \end{pmatrix} = \begin{pmatrix} 0 + x_1 \\ \vdots \\ 0 + x_n \end{pmatrix} .$$

Nach (2.2) und (2.4) lassen sich diese Gleichungen in der Form

$$x = x + 0 = 0 + x \tag{2.12}$$

schreiben. Die Gleichungen (2.12) bedeuten, dass der Nullvektor $0 \in \mathbb{R}^n$ das *neutrale Element* bezüglich der Addition von Spaltenvektoren ist. Aus der Rechenregel $-0 = 0$ für $0 \in \mathbb{R}$ folgt die entsprechende Regel

$$-0 = \begin{pmatrix} -0 \\ \vdots \\ -0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \tag{2.13}$$

für den Nullvektor $0 \in \mathbb{R}^n$.

Wir ziehen aus den Definitionen (2.1) und (2.2) weitere Konsequenzen. Aus der Assoziativität und der Kommutativität der Addition reeller Zahlen ergeben sich die entsprechenden Rechenregeln

$$(x + y) + z = x + (y + z)\,, \tag{2.14}$$
$$x + y = y + x \tag{2.15}$$

für die Addition von Spaltenvektoren. Das Assoziativgesetz für die Multiplikation reeller Zahlen liefert das Assoziativgesetz

$$(\alpha\beta)x = \alpha(\beta x) \tag{2.16}$$

für die Multiplikation von Skalaren und Spaltenvektoren. Die beiden Distributivgesetze für das Rechnen mit reellen Zahlen liefern die entsprechenden Rechengesetze

$$\alpha(x + y) = \alpha x + \alpha y\,, \tag{2.17}$$
$$(\alpha + \beta)x = \alpha x + \beta x \tag{2.18}$$

für Skalarmultiplikation und Spaltenvektoraddition.

Auf der Menge $\mathbb{R}_n$ der reellen Zeilen mit n-Komponenten werden durch

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)\,,$$

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

eine Multiplikation von reellen Zahlen mit Zeilen und eine Addition von Zeilen definiert. Siehe 2.10.

Wir fassen die obigen Ausführungen zusammen. Mit der Multiplikation (2.1) und der Addition (2.2) bilden die reellen Spaltenvektoren einen reellen Vektorraum im Sinne der folgenden Definition 2.2. Ein weiteres Beispiel liefern die reellen Polynome mit den üblichen Rechenregeln.

Definition 2.2 (Reelle Vektorräume)**.** *Ein Tripel* $(V, \cdot, +)$ *bestehend aus einer nicht-leeren Menge* V*, einer Abbildung* $\cdot : \mathbb{R} \times V \to V$ *mit*

$$(\lambda, x) \mapsto \lambda x = \lambda \cdot x$$

und einer Abbildung $+ \colon V \times V \to V$ *mit*

$$(x, y) \mapsto x + y$$

heißt ein reeller Vektorraum *oder ein* Vektorraum über $\mathbb{R}$*, wenn die folgenden Bedingungen* (1) *bis* (7) *erfüllt sind.*

(1) $(\forall x, y \in V) \colon x + y = y + x\,.$

(2) $(\forall x, y, z \in V) \colon (x + y) + z = x + (y + z)\,.$

(3) *Es gibt genau ein Element* $0 \in V$ *derart, dass* (3.1) *und* (3.2) *gelten.*

 (3.1) $(\forall x \in V) \colon x + 0 = 0 + x = x\,.$
 (3.2) $(\forall x \in V)(\exists! y \in V) \colon x + y = y + x = 0\,.$

(4) $(\forall x \in V) \colon 1x = 1 \cdot x = x\,.$

(5) $(\forall \alpha, \beta \in \mathbb{R})(\forall x \in V) \colon (\alpha\beta)x = \alpha(\beta x)\,.$

(6) $(\forall \alpha, \beta \in \mathbb{R})(\forall x \in V) \colon (\alpha + \beta)x = \alpha x + \beta x\,.$

(7) $(\forall \alpha \in \mathbb{R})(\forall x, y \in V) \colon \alpha(x + y) = \alpha x + \alpha y\,.$

Wir sagen dann auch, dass die Abbildungen $\cdot$ und $+$ auf V die Struktur eines reellen Vektorraumes definieren*.*

Um die Definition 2.2 nicht zu überfrachten, stellen wir die Erörterung einiger Bezeichnungen und Fachbegriffe nach. Wichtig sind die Rechengesetze (1) bis (7). Damit überhaupt etwas nennenswertes geschieht, verlangen wir, dass die *Grundmenge* V nicht leer sein möge.

Die Elemente der Grundmenge V des Vektorraumes $(V, \cdot, +)$ heißen *Vektoren*. Oft vereinfachen wir die Ausdrucksweise und sagen, dass V ein Vektorraum ist. Die Elemente von $\mathbb{R}$ heißen *Skalare*.

Die Abbildung $\cdot$ heißt die *Multiplikation mit Skalaren*. Das *Unitaritätsgesetz* (4) verhindert, dass diese Multiplikation trivial wird. Nach (5) ist die Multiplikation mit Skalaren *assoziativ*. Die Multiplikation mit Skalaren und die Multiplikation in $\mathbb{R}$ bezeichnen wir mit demselben Symbol.

Die *Vektoraddition* $+$ ist nach (1) *kommutativ* und nach (2) *assoziativ*. Sie besitzt nach (3.1) das neutrale *Element* 0. Der Vektor 0 heißt der *Nullvektor*.

Wie üblich bezeichnen wir die neutralen Elemente der Vektoradditon in V und der Addition in $\mathbb{R}$ mit demselben Symbol. Ebenso bezeichnen wir die Additionen in V und $\mathbb{R}$ mit demselben Symbol.

Sei $x \in V$. Nach (3.2) gibt es genau ein $y \in V$ mit $x + y = y + x = 0$. Dieser Vektor heißt der *negative Vektor zu* x und wird mit $-x$ oder $(-x)$ bezeichnet.

Die Multiplikation mit Skalaren und die Vektoraddition sind durch die beiden *Distributivgesetze* (6) und (7) miteinander verknüpft. In (6) und (7) verwenden wir die Konvention, dass eine Punktrechnung vor einer Strichrechnung ausgeführt wird.

Satz und Definition 2.3. *Sei* $(X, \cdot, +)$ *ein reeller Vektorraum. Eine nichtleere Teilmenge* $M \subseteq X$ *heißt ein* reeller Teilvektorraum *von* X*, wenn die beiden folgenden Bedingungen* (1) *und* (2) *erfüllt sind.*

(1) $(\forall \alpha \in \mathbb{R})(\forall x \in M): \quad \alpha x \in M.$

(2) $(\forall x, y \in M): \quad x + y \in M.$

Mit den Einschränkungen von $\cdot$ *und* $+$ *auf* $\mathbb{R} \times M$ *respektive* $M \times M$ *ist ein reeller Teilvektorraum* M *von* X *ein reeller Vektorraum.*

Definition 2.4 (Lineare Abbildungen)**.** *Seien* X, Y *reelle Vektorräume.*

(1) *Eine Abbildung* $\varphi \colon X \to Y$ *heißt* reell-linear *oder* linear*, wenn die beiden Bedingungen* (i) *und* (ii) *erfüllt sind.*

(i) $(\forall \alpha \in \mathbb{R})(\forall x \in X) \colon \varphi(\alpha x) = \alpha \varphi(x)\,.$

(ii) $(\forall x_1, x_2 \in X) \colon \varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2)\,.$

Die beiden Bedingungen (i) *und* (ii) *können durch die Bedingung* (iii) *gleichwertig ersetzt werden.*

(iii) $(\forall \alpha_1, \alpha_2 \in \mathbb{R})(\forall x_1, x_2 \in X) \colon \varphi(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \varphi(x_1) + \alpha_2 \varphi(x_2)\,.$

Wenn φ *linear ist, dann gilt* $\varphi(0) = 0$*. Lineare Abbildungen heißen auch* lineare Homomorphismen *oder* Homomorphismen.

(2) *Eine reell-lineare Abbildung* $\psi \colon X \to \mathbb{R}$ *heißt eine* reelle Linearform auf X*. Vereinfachend sagen wir, dass* ψ *eine Linearform ist. Siehe* 2.10.

(3) *Injektive Homomorphismes heißen* Monomorphismen*. Siehe* 23.3.

(4) *Surjektive Homomorphismes heißen* Epimorphismen*. Siehe* 23.3.

(5) *Eine bijektive lineare Abbildung* $\varphi \colon X \to Y$ *heißt ein* Isomorphismus *von* X *auf* Y*. Wir schreiben dann* $X \cong_\varphi Y$*. Die Umkehrabbildung* $\varphi^{-1} \colon Y \to X$ *ist ebenfalls ein Isomorphismus. Siehe* 23.3 *und* 23.5.

(6) *Wenn es einen Isomorphismus von* X *auf* Y *gibt, dann heißen* X *und* Y isomorph*. Wir schreiben dann* $X \cong Y$*. Isomorphe Vektorräume können miteinander identifiziert werden. Siehe* 2.10 *und* 2.21.

In 2.2, 2.3, 2.4 kann der Körper $\mathbb{R}$ der reellen Zahlen durch einen beliebigen Körper K im Sinne von Definition 15.1 ersetzt werden. Viele der folgenden Definitionen und Sätze machen auch über beliebigen Körpern Sinn. Wir gehen darauf mündlich in der Vorlesung ein. Im Abschnitt 16 arbeiten wir im Beweis der reellen Version 16.6 des Spektralsatzes vorübergehend in der Körpererweiterung $\mathbb{C}$ von $\mathbb{R}$.

Beispiele 2.5.

(1) Die *Nullabbildung* $\varphi : X \to Y$ mit $\varphi(x) = 0$ für alle $x \in X$ eine lineare Abbildung. Die Nullabbildung wird oft mit 0 bezeichnet.

(2) Die *identische Abbildung* $\mathrm{id}_X : X \to X$ mit $\mathrm{id}_X(x) = x$ für alle $x \in X$ ist eine lineare Abbildung.

(3) Für alle linearen Abbildungen $\varphi : X \to Y$ gilt $\varphi(0) = 0$.

(4) Die Menge $\mathrm{L}(X, Y)$ der linearen Abbildungen $\varphi : X \to Y$ ist mit

$$(\lambda\psi)(x) = \lambda\psi(x)\,, \quad (\varphi_1 + \varphi_2)(x) = \varphi_1(x) + \varphi_2(x)$$

ein reller Vektorraum.

Die folgenden Beispiele 2.6 und 2.7 bereiten spätere Überlegungen vor. In diesen Beispielen verwenden wir die Rechenoperationen (1.1), (1.2), (1.8) aus der Erörterung des einführenden Beispieles im ersten Abschnitt.

Beispiel 2.6 (Matrizen als lineare Abbildungen)**.** Sei $A = (a_{ij})_{i,j=1,2}$ eine reelle (2×2)-Matrix. Dann definiert A vermittels

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto Ax = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$

eine lineare Abbildung $\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2$. Siehe Satz 5.9. □

Beispiel 2.7 (Linearität der Determinante in jeder Spalte)**.** Seien $a = (a_1, a_2)^t$, $b = (b_1, b_2) \in \mathbb{R}^2$ gegeben. Dann definiert die Determinante vermittels

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \det \begin{pmatrix} x_1 & b_1 \\ x_2 & b_2 \end{pmatrix} = x_1 b_2 - x_2 b_1\,,$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \det \begin{pmatrix} a_1 & x_1 \\ a_2 & x_2 \end{pmatrix} = a_1 x_2 - a_2 x_1$$

zwei reelle Linearformen auf dem arithmetischen Vektorraum $\mathbb{R}^2$. *Dafür sagen wir auch, dass die Determinante in jeder Spalte linear ist.* Siehe Satz 12.5. □

Satz und Definition 2.8. *Seien X, Y reelle Vektorräume. Sei $\varphi\colon X \to Y$ eine lineare Abbildung.*

(1) *Der reelle Teilvektorraum*

$$\ker(\varphi) = \{\, x \in X \mid \varphi(x) = 0 \,\}$$

von X heißt der Kern *oder der* Nullraum *von φ.*

(2) *Der reelle Teilvektorraum*

$$\operatorname{im}(\varphi) = \{\, y \in Y \mid (\exists x \in X):\ \varphi(x) = y \,\}$$

von Y heißt das Bild *von φ.*

Zur Lösung eines linearen Gleichungssystems ist es wichtig, eine maximale Menge linear unabhängiger Spalten der Systemmatrix zu finden. Wir formulieren den Begriff der linearen Unabhängigkeit.

Satz und Definition 2.9 (Lineare Unabhängigkeit)**.** *Sei X ein reeller Vektorraum.*

(1) *Sei $m \in \mathbb{N}$. Eine endliche Familie $\mathcal{F}$ aus den Vektoren $x_1, \ldots, x_m \in X$ heißt* linear unabhängig über $\mathbb{R}$ *oder* linear unabhängig, *wenn aus*

$$\sum_{i=1}^{m} \alpha_i x_i \equiv \alpha_1 x_1 + \ldots + \alpha_n x_m = 0 \tag{2.19}$$

mit $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ stets

$$\alpha_1 = \ldots = \alpha_m = 0 \tag{2.20}$$

folgt. Andernfalls heißt die Familie $\mathcal{F}$ (über $\mathbb{R}$) linear abhängig .[1]

(2) *Die lineare Unabhängigkeit respektive die lineare Abhängigkeit ist eine Eigenschaft der Familie $\mathcal{F}$. Vereinfachend sagen wir, dass die Vektoren $x_1, \ldots, x_m \in X$ linear unabhängig respektive linear abhängig sind.*

(3) *Eine Summe der Form $\alpha_1 x_1 + \ldots + \alpha_m x_m$ heißt eine* endliche reelle Linearkombination *oder* Linearkombination *der Vektoren $x_1, \ldots, x_m$ mit den reellen* Koeffizienten $\alpha_1, \ldots, \alpha_m$. *Wenn alle Koeffizienten α_i verschwinden, dann heißt die Linearkombination* trivial.

(4) *Wenn mindestens zwei Vektoren der Familie $\mathcal{F}$ übereinstimmen, dann ist $\mathcal{F}$ linear abhängig.*

(5) *Sei $x \in X$ mit $x \neq 0$ gegeben. Dann ist die einelementige Menge x linear unabhängig.*

[1]Eine Familie ist eine Menge aus indizierten Elemente. Es kann durchaus sein, dass in einer Familie einige Elemente mehrfach vorkommen. Beispielsweise können mehrere Spalten einer Matrix übereinstimmen. Die Elemente einer Familie unterscheiden sich durch ihre Indizes. Jede Menge ist eine Familie.

(6) *Die leere Menge ist eine endliche Familie. Die leere Menge ist stets linear unabhängig.*

(7) *Eine beliebige Familie $\mathcal{G}$ aus Vektoren des Vektorraumes V heißt* linear unabhängig, *wenn jede endliche Teilfamilie von $\mathcal{G}$ linear unabhängig ist. Andernfalls heißt $\mathcal{G}$ linear abhängig. Wieder sagen wir vereinfachend, dass die Vektoren der Familie $\mathcal{G}$ linear unabhängig respektive linear abhängig sind.*

(8) *Sei $m \in \mathbb{N}$. Seien $x_1, \dots, x_m \in X$ linear unabhängig. Sei $x \in X$. Wenn die Vektoren $x_1, \dots, x_m, x$ linear abhänigig sind, dann ist der Vektor x eine Linearkombination der Vektoren $x_1, \dots, x_m \in X$.*

Beispiele 2.10. Siehe Beispiele 2.20.

(1) Die Teilmenge $\mathbb{R}^0 = \{0\} \subseteq \mathbb{R}$ ist mit

$$\lambda \cdot 0 = 0\,, \quad 0 + 0$$

ein reeller Vektorraum. Die leere Menge ist die einzige linear unabhängige Teilmenge von $\mathbb{R}^0$. Eine einelementige Menge $V = \{v\}$ ist mit

$$\lambda v = v\,, \quad v + v = v$$

ein reeller Vektorraumes. Dabei ist v das neutrale Element bezüglich der Vektoraddition. Alle reellen Vektoräume, die nur aus einem Element bestehen, können mit dem *kanonischen Nullraum* $\mathbb{R}^0$ identifiziert werden.

(2) Die reellen Zahlen bilden mit der üblichen Multiplikation und der üblichen Addition einen reellen Vektorraum $(\mathbb{R}, \cdot, +)$.

(3) Sei $n \in \mathbb{N}$. Dann bilden die n-komponentigen reellen Spaltenvektoren mit den komponentenweise erkärten Rechenoperationen (2.1)und (2.2) einen reellen Vektorraum $(\mathbb{R}^n, \cdot, +)$. Die n-komponentigen Spaltenvektoren

$$e_1^{(n)} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2^{(n)} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad, \quad e_n^{(n)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{2.21}$$

sind linear unabhängig. Wenn keine Verwechslungen zu befürchten sind schreiben wir e_i anstelle von $e_i^{(n)}$. Wir identifizieren $(\mathbb{R}, \cdot, +)$ mit $(\mathbb{R}^1, \cdot, +)$ vermittels $\lambda \mapsto (\lambda)$.

(4) Sei $n \in \mathbb{N}$. Wir formulieren die Definition des Vektorraumes $\mathbb{R}^n$ um. Dabei wird die spezielle Anordnung der Komponenten in einer Spalte zunächst vermieden. *Mit $\mathbb{R}^n$ bezeichnen wir die Menge der Abbildungen*

$$x \colon \{1, \dots, n\} \to \mathbb{R}\,. \tag{2.22}$$

Für $k = 1, \ldots, n$ setzen wir $x_k = x(k)$. Dann kann $x\colon \{1, \ldots, n\} \to \mathbb{R}$ mit dem Spaltenvektor

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

identifiziert werden. Auf dem Abbildungsraum $\mathbb{R}^n = \operatorname{Abb}(\{1, \ldots, n\}, \mathbb{R})$ definieren

$$(\lambda x)(k) = \lambda x(k)\,, \quad (x + y)(k) = x(k) + y(k) \tag{2.23}$$

die Struktur eines reellen Vektorraumes. Diese stimmt mit der durch (2.1) *und* (2.2) *definierten Struktur überein.*

(5) Die Linearformen auf $\mathbb{R}^n$ bilden mit

$$(\lambda\varphi)(x) = \lambda\varphi(x)\,, \quad (\varphi + \psi)(x) = \varphi(x) + \psi(x)$$

einen reellen Vektorraum $((\mathbb{R}^n)^*, \cdot, +)$, der mit dem Zeilenraum

$$\mathbb{R}_n = \{x^t \mid x \in \mathbb{R}^n\}$$

identifiziert werden kann. Siehe 7.4. Dabei wird eine Linearform $\varphi \in (\mathbb{R}^n)^*$ mit dem Zeilenvektor

$$(\varphi(e_1), \ldots, \varphi(e_n)) \in \mathbb{R}_n$$

identifiziert. Die Zeilen $e_1^t, \ldots, e_n^t$ sind linear unabhängig.

(6) Die reellen Polynome in einer Variablen x bilden mit den monomweise erklärten Rechenoperationen

$$\lambda \cdot \left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=0}^{n} \lambda a_i x^i\,, \tag{2.24}$$

$$\left(\sum_{i=0}^{n} a_i x^i\right) + \left(\sum_{i=0}^{n} b_i x^i\right) = \sum_{i=0}^{n} (a_i + b_i) x^i \tag{2.25}$$

einen reellen Vektorraum $(\mathbb{R}[x], \cdot, +)$. Die unendlich vielen Monome

$$x^0,\, x^1,\, x^2,\, x^3,\, \ldots \tag{2.26}$$

sind linear unabhängig.

Wir haben Spaltenvektoren und Matrizen eingeführt, um lineare Gleichungssysteme behandeln zu können. Umgekehrt führt auch das Rechnen mit Spaltenvektoren auf lineare Gleichungssysteme. Beispielsweise führt die Untersuchung auf lineare Unabhängigkeit auf ein homogenes Gleichungssystem.

Beispiel 2.11. Wir untersuchen die Vektoren

$$x_1 = \begin{pmatrix} 6 \\ 3 \end{pmatrix} , \quad x_2 = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

auf lineare Unabhängigkeit. Mit Satz 1.1 erhalten wir

$$\begin{aligned} \alpha_1 x_1 + \alpha_2 x_2 = 0 \quad &\Longleftrightarrow \quad \alpha_1 \begin{pmatrix} 6 \\ 3 \end{pmatrix} + \alpha_2 \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Longleftrightarrow \quad \begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Longleftrightarrow \quad \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} , \end{aligned}$$

denn es gilt $6 \cdot 5 - 3 \cdot 7 = 9 \neq 0$. Also sind die Spaltenvektoren x_1 und x_2 linear unabhängig. □

Beispiel 2.12. Die Vektoren

$$x_1 = \begin{pmatrix} 6 \\ 3 \end{pmatrix} , \quad x_2 = \begin{pmatrix} 24 \\ 12 \end{pmatrix}$$

sind linear abhängig. Offenbar gilt

$$(-4\alpha) \cdot \begin{pmatrix} 6 \\ 3 \end{pmatrix} + \alpha \cdot \begin{pmatrix} 24 \\ 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

für alle $\alpha \in \mathbb{R}$. Zum Nachweis der linearen Abhängigkeit von x_1 und x_2 wählen wir irgendein reelles $\alpha \neq 0$. □

Beispiel 2.13. Zwei beliebige Spaltenvektoren $a = (a_1, a_2)^t \in \mathbb{R}^2$ und $b = (b_1, b_2)^t \in \mathbb{R}^2$ sind genau dann linear abhängig, wenn

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = a_1 b_2 - a_2 b_1 = 0$$

gilt. Siehe Satz 13.1. □

Definition 2.14. *Sei X ein reeller Vektorraum.*

(1) *Sei $M \subseteq \mathbb{R}^n$ nicht-leer. Dann heißt die Menge $[M]$ aller $x \in X$ mit*

$$(\exists k \in \mathbb{N})(\exists \alpha_1, \ldots, \alpha_k \in \mathbb{R})(\exists a_1, \ldots a_k \in M) : \quad x = \sum_{\kappa=1}^{k} \alpha_\kappa a_\kappa$$

die lineare Hülle *von M. Die nicht-leere Menge $[M]$ ist ein Teilvektorraum von X.*

(2) *Wir betrachten den Nullraum $\mathbb{R}^0 = \{0\}$ als lineare Hülle $[\emptyset]$ der leeren Menge $\emptyset$.*

(3) *$M \subseteq X$ heißt ein* Erzeugendensystem *von X, wenn $[M] = X$ gilt. Wenn X ein endliches Erzeugendensystem besitzt, dann heißt X* endlich erzeugt.

(4) *$\mathcal{A} \subseteq X$ heißt eine* Basis *von X, wenn $\mathcal{A}$ ein Erzeudensystem von X ist und die Menge $\mathcal{A}$ linear unabhängig ist. Die leere Menge eine einzige Basis des Nullraumes.*

(5) *X heißt* endlich-dimensional, *wenn X eine endliche Basis besitzt.*

Satz 2.15. *Sei $X \neq \{0\}$ ein reeller Vektorraum. Eine Teilmenge*

$$\mathcal{A} = \{a_1, \dots, a_k\} \subseteq X$$

mit $k \in \mathbb{N}$ ist genau dann eine Basis von X, wenn die Bedingung

$$(\forall x \in X)(\exists! \alpha_1 \in \mathbb{R}) \dots (\exists! \alpha_k \in \mathbb{R}) : \quad x = \sum_{\kappa=1}^{k} \alpha_\kappa a_\kappa$$

erfüllt ist.

Wir zeigen im Folgenden, dass jeder endlich erzeugte reelle Vektorraum eine Basis besitzt. Dazu verwenden wir den Schrankensatz 2.17. Der Schrankensatz beruht auf dem Satz 2.16, nach dem ein homogenes lineares Gleichungssystem mit mehr Unbekannten als Gleichungen mindesten eine nicht-triviale Lösung besitzt. Dieser Satz wird durch vollständige Induktion bewiesen. In Satz 6.8, Aussage (4) formulieren wir diesen Satz in eine Aussage über den Kern bestimmter Matrizen um.

Satz 2.16. *Seien $m, n \in \mathbb{N}$ mit $m > n$ und $\alpha_{\nu\mu} \in \mathbb{R}$ für $\nu = 1, \dots, n$ und $\mu = 1, \dots, m$ gegeben. Ein homogenes lineares Gleichungssystem*

$$\sum_{\mu=1}^{m} \alpha_{\nu\mu} \beta_\mu = 0 \,, \quad \nu = 1, \dots, n$$

besitzt eine nicht-triviale Lösung $(\beta_1, \dots, \beta_m)^t \in \mathbb{R}^m$.

Beweis. Wir verwenden, dass die Lösungsmenge eines homogenen linearen Gleichungssystems durch elementare Zeilenumformungen nicht geändert wird.

Erster Schritt. Sei $m = n + 1$. In diesem Fall wird der Beweis vermittels vollständiger Induktion nach n erbracht.

Induktionsanfang. Für $n = 1$ erhalten wir die Gleichung

$$\alpha_{11}\beta_1 + \alpha_{12}\beta_2 = 0$$

in den Unbekannten β_1 und β_2. Diese Gleichung besitzt offenbar nicht-triviale Lösungen $(\beta_1, \beta_2)^t$.

Induktionsschluss. Wir betrachten

$$\sum_{\mu=1}^{n+2} \alpha_{\nu\mu}\beta_\mu = 0\,, \quad \nu = 1, \ldots, n+1\,.$$

Dabei können wir annehmen, dass die Systemmatrix $(\alpha_{\nu\mu})$ in oberer Stufenform vorliegt. Diese Situtation kann durch elementare Zeilenumformungen erreicht werden. Wir nehmen eine Fallunterscheidung von. Die folgenden drei Fälle können auftreten.

(1) Die $(n+1)$-te Zeile der Systemmatrix sei die Nullzeile. Dann setzen wir $\beta_{n+2} = 0$. Nach Induktionsvoraussetzung gibt es $\beta_1, \ldots, \beta_{n+1} \in \mathbb{R}$ derart, dass $(\beta_1, \ldots, \beta_{n+1}, 0)^t$ eine nicht-triviale Lösung ist.

(2) Sei α_{n+1n+1} der erste Koeffizient, der in der $(n+1)$-ten Zeile nicht verschwindet. In diesem Fall können wir annehmen, dass sämtliche Koeffizienten in der $(n+1)$-Spalte, die oberhalb des Diagonalelementes α_{n+1n+1} liegen, verschwinden. Nach der Induktionsvoraussetzung gibt es β_μ mit $\mu \neq n+1$, die eine nicht-triviale Lösung der ersten n Gleichungen bilden. Dann wird β_{n+1} so bestimmt, dass auch die $(n+1)$-te Gleichung erfüllt ist. Damit ist eine nicht-triviale Lösung $(\beta_1, \ldots, \beta_{n+2})^t$ gefunden.

(3) Sei α_{n+1n+2} der erste Koeffizient, der in der $(n+1)$-ten Zeile nicht verschwindet. Wir setzen $\beta_{n+2} = 0$. Nach Induktionsvoraussetzung gibt es $\beta_1, \ldots, \beta_{n+1} \in \mathbb{R}$ derart, dass $(\beta_1, \ldots, \beta_{n+1}, 0)^t$ eine nicht-triviale Lösung ist.

Damit ist der Induktionsschluss beendet.

Zweiter Schritt. Sei $m > n+1$. Sei $(\beta_1, \ldots, \beta_{n+1})^t$ eine nicht-triviale Lösung des Gleichungssystems

$$\sum_{\mu=1}^{n+1} \alpha_{\nu\mu}\beta_\mu = 0\,, \quad \nu = 1, \ldots, n\,.$$

Wir setzen $\beta_{n+2} = \ldots = \beta_m = 0$. Dann ist $(\beta_1, \ldots, \beta_m)^t$ eine nicht-triviale Lösung von

$$\sum_{\mu=1}^{m} \alpha_{\nu\mu}\beta_\mu = 0\,, \quad \nu = 1, \ldots, n\,.$$

Damit ist der Beweis beendet. □

Satz 2.17 (Schrankensatz)**.** *Sei X ein reeller Vektorraum. Sei $n \in \mathbb{N}_0$. Wenn X ein Erzeugendensysten aus n Elementen besitzt, dann sind $n+1$ Vektoren aus X linear abhängig.*

Beweis. Der Satz gilt offenbar im Fall $X = \{0\}$. Wir setzen $X \neq \{0\}$ voraus. Die Menge $\{a_1, \dots a_n\} \subseteq X$ mit $n \in \mathbb{N}$ sei ein Erzeugendensystem von X. Seien $x_1, \dots, x_{n+1} \in X$ beliebig gewählt. Dann gibt es $\alpha_{\nu\mu} \in \mathbb{R}$ mit

$$x_\mu = \sum_{\nu=1}^{n} \alpha_{\nu\mu} a_\nu \,, \quad \mu = 1, \dots, n+1 \,. \tag{2.27}$$

Wir betrachten eine Linearkombination

$$\sum_{\mu=1}^{n+1} \beta_\mu x_\mu \,.$$

Einsetzen von (2.27) ergibt

$$\sum_{\mu=1}^{n+1} \beta_\mu x_\mu = \sum_{\nu=1}^{n} \left(\sum_{\mu=1}^{n+1} \alpha_{\nu\mu} \beta_\mu \right) a_\nu \,.$$

Wenn das homogene lineare Gleichungssystem

$$\sum_{\mu=1}^{n+1} \alpha_{\nu\mu} \beta_\mu = 0 \,, \quad \nu = 1, \dots, n \tag{2.28}$$

aus n Gleichungen und in den Unbekannten $\beta_1, \dots, \beta_{n+1}$ eine nicht-triviale Lösung besitzt, dann sind $x_1, \dots, x_{n+1}$ linear abhängig. Nach Satz 2.16 gibt es eine nicht-triale Lösung von 2.28. Also sind $x_1, \dots, x_{n+1}$ linear abhängig. □

Satz 2.18. *Sei $X \neq \{0\}$ ein endlich erzeugte reeller Vektorraum.*

(1) *Sei $\mathcal{A} = \{a_1, \dots, a_r\} \subseteq X$ mit $r \in \mathbb{N}$ linear unabhängig. Dann ist $\mathcal{A}$ eine Basis von X oder es gibt $a_{r+1}, \dots, a_m \in X$ mit $m \in \mathbb{N}$ und $m > r$ derart, dass $\mathcal{C} = \{a_1, \dots, a_m\}$ eine Basis von X ist.*

(2) *X besitzt eine endliche Basis.*

(3) *Zwei Basen von X besitzen gleich viele Elemente.*

Beweis. Wir wenden den Schrankensatz 2.17 mehrfach an.

Nachweis von (1). Die linear unabhängige Menge $\mathcal{A} = \{a_1, \dots, a_r\}$ sei keine Basis von X. Dann gibt es $a_{r+1} \in X$ mit $a_{r+1} \notin [\mathcal{A}]$. Die erweiterte Menge

$$\mathcal{A}_1 = \mathcal{A} \cup \{a_{r+1}\}$$

ist linear unabhängig. Dieses Verfahren der Hinzunahme eines weiteren Vektors derart, dass die erweiterte Menge ebenfalls linear unabhängig ist, bricht nach endlich vielen Schritten ab. Dies folgt aus dem Schrankensatz 2.17, weil X endlich erzeugt ist. Also gibt es $m \in \mathbb{N}$ derart, dass $\mathcal{A}_m$ eine linear unabhängige Menge mit $|\mathcal{A}_m| = X$ ist. Die Menge $\mathcal{A}_m$ wird im $(m-r)$-ten Schritt erreicht.

Nachweis von (2). Wegen $X \neq \{0\}$ gibt es $v_1 \in X \setminus \{0\}$. Also ist $\{v_1\}$ linear unabhängig. Wenn $\{v_1\}$ noch keine Basis von X ist, dann kann $\{v_1\}$ nach (1) zu einer endlichen Basis von X ergänzt werden.

Nachweis von (3). Mit $|M|$ bezeichnen wir die Anzahl der Elemente einer endlichen Menge. Sei $\mathcal{A}$ eine endliche Basis von X. Nach dem Schrankensatz 2.17 sind $|\mathcal{A}| + 1$ Elemente von X linear abhängig. Also ist jede Basis von X endlich. Außerdem gilt $|\mathcal{B}| \leq |\mathcal{A}|$ für irgendeine andere Basis $\mathcal{B}$ von X. Wir beginnen mit der Basis $\mathcal{B}$. Nach dem Schrankensatz 2.17 gilt $|\mathcal{A}| \leq |\mathcal{B}|$. Damit ist auch (3) bewiesen. □

Der Nullraum ist endlich erzeugt. Er besitzt nur die leere Menge als Basis. Es gilt $|\emptyset| = 0$. Die leere Menge ist endlich. Wegen Satz 2.18 ist daher die folgende Begriffsbildung 2.19 sinnvoll. Ein reeller Vektorraum ist genau dann endlich-dimensional, wenn er endlich erzeugt ist. Beliebige reelle Vektorräume besitzen ebenfalls Basen. Dabei sind zwei Basen gleichmächtig. Den Dimensionsbegriff formulieren wir in dieser Vorlesung nur im Fall endlich erzeugter reeller Vektorräume.

Definition 2.19.

(1) *Sei X ein endlich erzeugter reeller Vektorraum. Die Anzahl $|\mathcal{A}|$ der Elemente einer beliebigen Basis $\mathcal{A}$ von X heißt die* Dimension *von X. Die Dimension von X hängt nicht von der Wahl der Basis ab. Die Dimension von X bezeichnen wir mit $\dim(X)$. Offenbar gilt $\dim(X) \in \mathbb{N}_0$.*

(2) *Sei X ein reeller Vektorraum, der nicht endlich erzeugt ist. Dann setzen wir $\dim(X) = \infty$ und sagen, dass Y* unendlich-dimensional *ist.*

(3) *Wir schreiben $\dim(X) < \infty$ genau dann, wenn X endlich-dimensional ist.*

Beispiele 2.20. Siehe Beispiele 2.10.

(1) $\dim(\mathbb{R}^0) = 0$.

(2) Sei $n \in \mathbb{R}^n$. Die Menge $\mathcal{E}_n = \{e_1^{(n)}, \ldots, e_n^{(n)}\}$ ist eine Basis von $\mathbb{R}^n$. Die Basis $\mathcal{E}_n$ heißt die *kanonische Basis des* $\mathbb{R}^n$. Es gilt $\dim(\mathbb{R}^n) = n$.

(3) Die Monome x^n mit $n \in \mathbb{N}_0$ bilden eine Basis des Vektoraumes $\mathbb{R}[x]$ der reellen Polynome in der Unbestimmten x. Diese Basis ist nicht endlich.

Ein reeller Vektorraum X mit $\dim(X) < \infty$ kann mit dem Vektorraum $\mathbb{R}^n$, wobei $n = \dim(X)$ gilt, identifiziert werden.

Satz 2.21. *Sei X ein reeller Vektorraum mit $n = \dim(X) < \infty$.*

(1) *Sei $n = 0$. Dann ist die Nullabbildung 0 der einzige Isomorphimus von X auf $\mathbb{R}^0$. Es gilt also $X \cong_0 \mathbb{R}^0$.*

(2) *Sei $n \in \mathbb{N}$. Dann gibt es zu jeder Basis $\mathcal{A} = \{a_1, \dots, a_n\}$ genau einen Isomorphismus $\varphi\colon X \to \mathbb{R}^n$ mit $\varphi(a_i) = e_i^{(n)}$ für alle $i = 1, \dots, n$. Es gilt also $X \cong_\varphi \mathbb{R}^n$.*

(3) *$X \cong \mathbb{R}^n$.*

(4) *Sei $Y \subseteq X$ ein reeller Teilvektorraum von X. Dann gelten:*

   (i) $\dim(Y) \leq \dim(X)$.

   (ii) $Y = X \Leftrightarrow \dim(Y) = \dim(X)$

3 Inneres Produkt. Euklidische Vektorräume

In diesem Abschnitt definieren wir das *euklidische innere Produkt* zweier Spaltenvektoren des $\mathbb{R}^n$. Wenn keine Verwechslungen zu befürchten sind, sprechen wir einfach vom inneren Produkt. Für das innere Produkt ist auch der Name *Skalarprodukt* üblich. Wir ziehen die Bezeichnung inneres Produkt vor, um Verwechslungen mit der *Skalarmultiplikation* (2.1) zu vermeiden. Das innere Produkt gestattet es wiederum, Länge von Vektoren, den eingeschlossenen Winkel zwischen zwei Vektoren und den wichtigen Begriff der Orthogonalität algebraisch zu definieren. Mit diesem Instrumentarium können Begiffe und Resultate der Elementargeometrie einer algebraischen Behandlung zugänglich gemacht werden. Umgekehrt ermöglichen es diese geometrischen Anwendungen, Begriffe und Resulatate der linearen Algebra geometrisch zu deuten.

Definition 3.1. *Das* euklidische innere Produkt *oder einfach* innere Produkt *zweier Spaltenvektoren* $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ *und* $y = (y_1, \dots, y_n)^t \in \mathbb{R}^n$ *wird durch*

$$\langle x, y\rangle \equiv x \cdot y \equiv \sum_{i=1}^{n} x_i y_i \equiv x_1 y_1 + \ldots + x_n y_n \in \mathbb{R} \tag{3.1}$$

definiert.

Auf Grund der Schreibweise $x \cdot y$ mit einem fett gedruckten Punkt als Multiplikationszeichen wird das innere Produkt in der angelsächsischen Literatur auch als *dot product* bezeichnet. Wir verwenden im Folgenden beide Schreibweisen. Wenn die Spalten ausgeschrieben werden, empfiehlt sich der Lesbarkeit halber die Schreibweise mit einem fett gedruckten Punkt.

Beispiel 3.2. Für die Spaltenvektoren x, $y \in \mathbb{R}^4$ mit

$$x = \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix}, \quad y = \begin{pmatrix} 45 \\ 0 \\ -21 \\ 14 \end{pmatrix}$$

erhalten wir

$$\begin{aligned} \langle x, y\rangle &= \langle \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix}, \begin{pmatrix} 45 \\ 0 \\ -21 \\ 14 \end{pmatrix} \rangle \\ &= (-3) \cdot 45 + 7 \cdot 0 + 23 \cdot (-21) + (-25) \cdot 14 \\ &= -135 + 0 - 483 - 350 = -968\,. \end{aligned}$$

In der Schreibweise mit dem fett gedruckten Punkt gestaltet sich die erste Zeile etwas übersichtlicher. Am Ergebnis ändert sich natürlich nichts.

$$\begin{aligned} x \boldsymbol{\cdot} y &= \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix} \boldsymbol{\cdot} \begin{pmatrix} 45 \\ 0 \\ -21 \\ 14 \end{pmatrix} \\ &= (-3) \cdot 45 + 7 \cdot 0 + 23 \cdot (-21) + (-25) \cdot 14 \\ &= -135 + 0 - 483 - 350 = -968\,. \end{aligned}$$

Manchmal mischen wir auch die Schreibweisen.

$$\begin{aligned} \langle x, x \rangle &= \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix} \boldsymbol{\cdot} \begin{pmatrix} -3 \\ 7 \\ 23 \\ -25 \end{pmatrix} \\ &= (-3)^2 + 7^2 + 23^2 + (-25)^2 \\ &= 9 + 49 + 529 + 625 = 1212\,. \end{aligned}$$

□

Beispiel 3.3. Für die Spaltenvektoren

$$x_1 = \begin{pmatrix} 6 \\ 3 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

erhalten wir

$$\begin{aligned} \langle x_1, x_1 \rangle &= 36 + 9 = 45\,, \\ \langle x_1, x_2 \rangle &= 42 + 15 = 57\,, \\ \langle x_2, x_2 \rangle &= 49 + 25 = 74\,. \end{aligned}$$

Folglich gilt

$$\begin{aligned} \langle x_1, x_2 \rangle \langle x_1, x_2 \rangle &= 57 \cdot 57 \\ &= 3249 \\ &< 3330 \\ &= 45 \cdot 74 \\ &= \langle x_1, x_1 \rangle \langle x_2, x_2 \rangle\,. \end{aligned}$$

Hier haben wir ein Beispiel für die *Cauchy-Schwarz'sche Ungleichung* kennengelernt. □

Wir ziehen erste Konsequenzen aus der definierenden Gleichung (3.1) und formulieren die grundlegenden Rechengesetze für innere Produkte. Diese Gesetze behalten wir bei, wenn wir den Begriff des inneren Produktes vom Modell der Menge der Spaltenvektoren auf beliebige reelle Vektorräume übertragen.

Satz 3.4. *Seien $x,\ y,\ z \in \mathbb{R}^n$ und $\alpha,\ \beta \in \mathbb{R}$ beliebig gegeben. Dann gelten die folgenden sechs Aussagen.*

(1) $\langle x, y\rangle = \langle y, x\rangle$.

(2) $\langle x, x\rangle \geq 0$.

(3) $\langle x, x\rangle = 0 \Longleftrightarrow x = 0$.

(4) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$.

(5) $\langle z, \alpha x + \beta y\rangle = \alpha\langle z, x\rangle + \beta\langle z, y\rangle$.

(6) $\langle x, 0\rangle = \langle 0, x\rangle = 0$.

Beweis. Die Aussagen (1) bis (6) folgen direkt aus der Definition (3.1). Wir überlassen die einfachen Rechnungen dem geneigten Auditorium zur Übung als Hausaufgabe. □

Wir führen einige Fachbegriffe ein, die die Eigenschaften (1) bis (5) innerer Produkte benennen. Nach (1) ist das innere Produkt *symmetrisch*. Nach (2) und (3) ist es *positiv definit*. Nach (4) ist das innere Produkt bei festgehaltenem zweiten Argument bezüglich des ersten Argumentes *linear*. Nach (5) ist das innere Produkt bei festgehaltenem ersten Argument bezüglich des zweiten Argumentes *linear*. Beide Eigenschaften (4) und (5) fassen wir zusammen, indem wir sagen, dass das innere Produkt *bilinear* ist. Eigenschaft (6) folgt aus der Bilinearität. Der Vektorraum $\mathbb{R}^n$ ist mit dem euklidischen inneren Produkt ein *euklidischer Vektorraum* im Sinne der folgenden Definition.

Definition 3.5. *Sei X ein reeller Vektorraum mit* $\dim(X) < \infty$. *Eine Abbildung $\langle\,.\,,\,.\,\rangle\colon X \times X \to \mathbb{R}$, die die Eigenschaften* (1) *bis* (5) *aus Satz* 3.4 *für alle $x,\ y,\ z \in X$ und alle $\alpha,\ \beta \in \mathbb{R}$ erfüllt, heißt ein* inneres Produkt *auf X. Das Paar $(X, \langle\,.\,,\,.\,\rangle)$ heißt dann ein* euklidischer Vektorraum.

Viele der Sätze, die wir im Folgenden für den $\mathbb{R}^n$ und das euklidische innere Produkt formulieren, gelten auch in beliebigen euklidischen Vektorräumen. Die Beweise sind so formuliert, dass sie sich meist direkt auf den allgemeinen Fall übertragen lassen.

Satz 3.6 (Cauchy-Schwarz'sche Ungleichung)**.** *Seien $x,\ y \in \mathbb{R}^n$ zwei beliebige Spaltenvektoren. Dann gilt*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle\,. \tag{3.2}$$

Das Gleichheitszeichen gilt genau dann, wenn die beiden Spaltenvektoren x und y linear abhängig sind.

Bemerkungen. Ungleichung (3.2) heißt *Cauchy-Schwarz'sche Ungleichung.* Mit ihrer Hilfe werden wir die Dreiecks-Ungleichung beweisen. Vergleiche Satz 3.10, Aussage (4). Die Betragsstriche auf der linken Seite der Ungleichung sind beim Rechnen über $\mathbb{R}$ überflüssig. Auf die Betragsstriche kann in der analogen Formel für komplexe Spaltenvektoren allerdings nicht verzichtet werden. Wenn wir auf beiden Seiten die positive Quadratwurzel ziehen, dann müssen die Betragsstriche auch beim Rechnen über $\mathbb{R}$ stehen bleiben. Vergleiche (3.6).

Beweis. Erster Schritt. Im Fall $y = 0$ ist die Beziehung (3.2) mit dem Gleichheitszeichen erfüllt. Außerdem sind x und y in diesem Fall linear abhängig.

Zweiter Schritt. Wir nehmen im Folgenden an, dass $y \neq 0$ gilt. Sei $\lambda \in \mathbb{R}$ beliebig gewählt. Dann gilt

$$\begin{aligned} 0 &\leq \langle x - \lambda y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \lambda \langle x, y \rangle - \lambda \langle y, x \rangle + (-\lambda)^2 \langle y, y \rangle \\ &= \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle \,. \end{aligned}$$

Wir treffen eine spezielle Wahl für λ. Sei

$$\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle} \,.$$

Dieser Quotient ist sinnvoll gebildet, weil wir $y \neq 0$ vorausgesetzt haben. Einsetzen ergibt

$$0 \leq \langle x, x \rangle - 2 \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle x, y \rangle + \frac{\langle x, y \rangle^2}{\langle y, y \rangle^2} \langle y, y \rangle \,.$$

Nach der Multiplikation beider Seiten mit $\langle y, y \rangle > 0$ erhalten wir

$$0 \leq \langle x, x \rangle \langle y, y \rangle - 2 \langle x, y \rangle^2 + \langle x, y \rangle^2 = \langle x, x \rangle \langle y, y \rangle - \langle x, y \rangle^2 \,.$$

Nun folgt die behauptete Ungleichung

$$|\langle x, y \rangle|^2 = \langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle \,.$$

Dritter Schritt. Wir untersuchen im Fall $y \neq 0$, wann in (3.2) das Gleichheitszeichen gilt. Wenn x und y linear abhängig sind, dann gibt es $\mu \in \mathbb{R}$ mit $x = \mu y$. Wir erhalten

$$|\langle x, y \rangle|^2 = \langle \mu y, y \rangle^2 = \mu^2 \langle y, y \rangle^2 = \langle \mu y, \mu y \rangle \langle y, y \rangle = \langle x, x \rangle \langle y, y \rangle \,.$$

Damit folgt die Gleichheit in (3.2). Nun setzen wir umgekehrt die Gleichheit voraus. Dann folgt

$$\begin{aligned} \left\langle x - \frac{\langle x, y \rangle}{\langle y, y \rangle}\, y \,,\, x - \frac{\langle x, y \rangle}{\langle y, y \rangle}\, y \right\rangle &= \langle x, x \rangle - 2\, \frac{\langle x, y \rangle^2}{\langle y, y \rangle} + \frac{\langle x, y \rangle^2}{\langle y, y \rangle} \\ &= \frac{\langle x, x \rangle \langle y, y \rangle - \langle x, y \rangle^2}{\langle y, y \rangle} = 0 \,. \end{aligned}$$

Mit Aussage (3) des Satzes 3.4 folgt schließlich

$$x - \frac{\langle x, y\rangle}{\langle y, y\rangle}\, y = 0\,.$$

Folglich sind x und y linear abhängig. □

Definition 3.7 (Euklidische Norm)**.** *Sei $n \in \mathbb{N}$.*

(1) *Die* euklidische Länge *oder* euklidische Norm *eines Spaltenvektors*

$$x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$$

wird durch

$$\|x\| \equiv \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \ldots + x_n^2} \geq 0 \tag{3.3}$$

definiert. Wenn keine Verwechslungen zu befürchten sind, sprechen wir einfach von der Länge oder der Norm eines Spaltenvektors.

(2) *Ein Spaltenvektor $u \in \mathbb{R}^n$ heißt* normierter Vektor, Einheitsvektor *oder* Richtungsvektor*, wenn*

$$\|u\| = 1 \tag{3.4}$$

gilt. Einheitsvektoren definieren Richtungen*.*

(3) *Wenn $x \in \mathbb{R}^n$ nicht der Nullvektor ist, dann ist*

$$\frac{x}{\|x\|} \equiv \frac{1}{\|x\|}\, x \tag{3.5}$$

ein normierter Vektor. Den Einheitsvektor (3.5) *nennen wie die* Richtung des Vektors $x \neq 0$*. Dem Nullvektor ordnen wir keine Richtung zu.*

(4) *Die Einheitsvektoren des $\mathbb{R}^n$ bilden die* Einheitssphäre S^{n-1}*.*

Beispiele 3.8.

(1) *Wie üblich identifizieren wir die reellen Vektorräume $\mathbb{R}^1$ und $\mathbb{R}$ vermittels der Abbildung $(\alpha) \mapsto \alpha$. Dann gilt $S^0 = \{-1, 1\} \subseteq \mathbb{R}$.*

(2) *Die Sphäre $S^1 \subseteq \mathbb{R}^2$ ist die Einheitskreislinie in der Ebene um den Ursprung. Sie besteht aus allen Punkten der Form*

$$\begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$$

mit $\varphi \in [0, 2\pi)$. Die Winkelfunktionen Cosinus *und* Sinus *sind durch Reihenentwicklungen für alle $x \in \mathbb{R}$ definiert. Es gelten*

$$\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\, x^{2k}\,, \quad \sin(x) = \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!}\, x^{2l+1}\,.$$

Diese Reihen konvergieren in allen $x \in \mathbb{R}$ absolut. Siehe Analysis-Skript, Abschnitt 5*, Satz und Definition* 5.1*.*

Nach Aussage (2) von Satz 3.4 kann die positive Quadratwurzel aus $\langle x, x\rangle$ gezogen werden. Mit Hilfe der Norm kann die Cauchy-Schwarz'sche Ungleichung (3.2) wie folgt geschrieben werden.

Satz 3.9 (Cauchy-Schwarz'sche Ungleichung)**.** *Seien x, $y \in \mathbb{R}^n$ zwei beliebige Spaltenvektoren. Dann gilt*

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| \,. \tag{3.6}$$

Das Gleichheitszeichen gilt genau dann, wenn die beiden Spaltenvektoren x und y linear abhängig sind.

Satz 3.10. *Seien x, $y \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$ beliebig gegeben. Dann gelten die folgenden sechs Aussagen.*

(1) $\|x\| \geq 0$.

(2) $\|x\| = 0 \iff x = 0$.

(3) $\|\alpha x\| = |\alpha| \, \|x\|$.

(4) $\| - x\| = \|(-1)x\| = \|x\|$.

(5) $\|x + y\| \leq \|x\| + \|y\|$.

Das Gleichheitszeichen gilt in Ungleichung (5) *genau dann, wenn $y = 0$ oder $x = \lambda y$ mit $\lambda \geq 0$ erfüllt ist.*

(6) $| \, \|x\| - \|y\| \, | \leq \|x - y\|$.

Das Gleichheitszeichen gilt in Ungleichung (6) *genau dann, wenn $y = 0$ oder $x = \lambda y$ mit $\lambda \geq 0$ erfüllt ist.*

Bemerkungen. Die Ungleichung (5) des Satzes 3.10 heißt Dreiecks-Ungleichung. Auf die Bedingung $\lambda \geq 0$ kann im Zusatz zu Ungleichung (5) nicht verzichtet werden, wie das Beispiel $x = -y \neq 0$ zeigt.

$$\begin{aligned} 0 &= \|0\| = \|x + (-x)\| \,, \\ 0 &< 2\|x\| = \|x\| + \|x\| = \|x\| + \| - x\| \,. \end{aligned}$$

Auf der linken Seite der Ungleichung (6) steht der Betrag einer Differenz zweier nicht-negativer reeller Zahlen. Auf die Bedingung $\lambda \geq 0$ kann auch im Zusatz zu Ungleichung (6) nicht verzichtet werden, wie wieder das Beispiel $x = -y \neq 0$ zeigt.

$$\begin{aligned} 0 &= |0| = | \, \|x\| - \|x\| \, | = | \, \|x\| - \| - x\| \, | \,, \\ 0 &< 2\|x\| = \|x + x\| = \|x - (-x)\| \,. \end{aligned}$$

Wir wenden uns dem Beweis des Satzes 3.10 zu.

Beweis. Die Aussagen (1) bis (3) folgen nach leichter Rechnung aus der Definition (3.3). Aussage (4) folgt aus (3) mit $\alpha = -1$. Der Nachweis der Dreiecks-Ungleichung (5) ist auf direktem Wege beschwerlich. Wir folgern (5) mit Hilfe der Cauchy-Schwarz'sche Ungleichung.

Wir beginnen mit dem Quadrat der linken Seite der Dreiecks-Ungleichung. Nach Definition (3.3) und der Cauchy-Schwarz'sche Ungleichung (3.6) gilt

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \\
&\le \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2\,.
\end{aligned}$$

Die behauptete Ungleichung (5) folgt nun durch Übergang zu den positiven Quadratwurzeln.

Wir beweisen den Zusatz zu Ungleichung (5). Im Fall $y = 0$ gilt das Gleichheitszeichen in (5). Wir setzen $y \neq 0$ voraus. Die vorstehende Rechnung und der Zusatz zur Cauchy-Schwarz'schen Ungleichung in Satz 3.9 zeigen die Gültigkeit der Implikationen

$$\begin{aligned}
\|x+y\| = \|x\| + \|y\| \quad &\Longrightarrow \quad \langle x,y\rangle = \|x\|\,\|y\| \\
&\Longrightarrow \quad (\exists \lambda \in \mathbb{R}): \quad x = \lambda y\,.
\end{aligned}$$

Einsetzen von $x = \lambda y$ in $\langle x,y\rangle = \|x\|\,\|y\|$ ergibt $\lambda \ge 0$. Umgekehrt folgt für $x = \lambda y$ mit $\lambda \ge 0$ die Gültigkeit von

$$\begin{aligned}
\|x+y\| &= \|\lambda y + y\| \\
&= (\lambda + 1)\,\|y\| \\
&= \lambda\,\|y\| + \|y\| \\
&= \|\lambda y\| + \|y\| = \|x\| + \|y\|\,.
\end{aligned}$$

Damit ist der Zusatz zu Ungleichung (5) bewiesen.

Wir beweisen Ungleichung (6). Dazu wenden wir die Dreiecks-Ungleichung zweimal an. Aus den beiden Ungleichungen

$$\begin{aligned}
\|x\| &= \|(x-y) + y\| \le \|x-y\| + \|y\|\,, \\
\|y\| &= \|x + (y-x)\| \le \|x\| + \|y-x\| = \|x\| + \|x-y\|
\end{aligned}$$

folgen durch Umordnen die beiden Ungleichungen

$$\|x\| - \|y\| \le \|x-y\|\,,$$

$$\|y\| - \|x\| \leq \|x - y\|\,.$$

Zusammenfassen ergibt die behauptete Ungleichung

$$|\,\|x\| - \|y\|\,| \leq \|x - y\|\,.$$

Damit ist Ungleichung (6) bewiesen.

Wir beweisen den Zusatz zu Ungleichung (6). Es ist klar, dass im Fall $y = 0$ Gleichheit gilt. Für $y \neq 0$ und $x = \lambda y$ mit $\lambda \geq 0$ gilt

$$|\,\|x\| - \|y\|\,| = |\,\lambda \cdot \|y\| - 1 \cdot \|y\|\,| = |\lambda - 1|\,\|y\| = \|(\lambda - 1)y\| = \|x - y\|\,.$$

Nun setzen wir $y \neq 0$ und $|\,\|x\| - \|y\|\,| = \|x - y\|$ voraus. Zuerst betrachten wir den Fall $\|x\| \geq \|y\|$. Es folgt

$$\|x\| = \|x - y\| + \|y\|\,.$$

Nach dem Zusatz zur Dreiecks-Ungleichung gibt es ein $\alpha \geq 0$ mit

$$x - y = \alpha y\,, \quad x = (1 + \alpha)y\,.$$

Also liefert $\lambda = 1 + \alpha \geq 0$ die behauptete Darstellung der Form $x = \lambda y$ mit nicht-negativem λ. Abschließend betrachten wir den Fall $\|y\| > \|x\|$. Es folgt

$$\|y\| = \|x - y\| + \|x\| = \|x\| + \|y - x\|\,, \quad y - x \neq 0\,.$$

Nach dem Zusatz zur Dreiecks-Ungleichung gibt es ein $\beta \geq 0$ mit

$$x = \beta(y - x)\,, \quad x = \tfrac{\beta}{1+\beta}\,y\,.$$

Also liefert $\lambda = \frac{\beta}{1+\beta} \geq 0$ die behauptete Darstellung der Form $x = \lambda y$ mit nicht-negativem λ. Damit ist der Zusatz zu Ungleichung (6) bewiesen und der Beweis des Satzes 3.10 beendet. □

Definition 3.11. *Seien x, $y \in \mathbb{R}^n$ vom Nullvektor verschiedene Spaltenvektoren. Der* Cosinus des von x und y eingeschlossenen Winkels *wird durch*

$$\cos(x, y) \equiv \frac{\langle x, y\rangle}{\|x\|\,\|y\|} \tag{3.7}$$

definiert.

Diese Definition macht nur Sinn, wenn die Vektoren x und y normierbar sind. Wegen der Cauchy-Schwarz'sche Ungleichung (3.6) gilt

$$-1 \leq \cos(x, y) \leq 1\,. \tag{3.8}$$

Also macht es Sinn, die rechte Seite von (3.7) als Cosinus einer eindeutig bestimmten reellen Zahl $\varphi \in [0, \pi]$ aufzufassen, wobei φ geometrisch als *unorientierter eingeschlossener Winkel* gedeutet wird. Wir schreiben $\measuredangle(x, y)$ für φ. Damit erhalten wir

$$\cos(\measuredangle(x, y)) = \cos(x, y) \equiv \frac{\langle x, y\rangle}{\|x\|\,\|y\|}\,.$$

Wir fassen zusammen:

Satz und Definition 3.12 (Unorientierter Winkel)**.** Seien $x,\ y \in \mathbb{R}^n$ vom Nullvektor verschiedene Spaltenvektoren.

(1) Die reelle Zahl $\measuredangle(x, y) \in [0, \pi]$ mit

$$\measuredangle(x, y) = \arccos\left(\frac{\langle x, y\rangle}{\|x\|\,\|y\|}\right)$$

heißt der *unorierentierte Winkel zwischen x und y*. Dabei ist

$$\arccos\colon [-1, 1] \to [0, \pi] \subseteq \mathbb{R}$$

die Umkehrfunktion der Einschränkung des Cosinus auf das Intervall $[0, \pi]$.

(2) In der Analysis werden Funktionen oft durch Integrale definiert. Es gilt

$$\arccos(\tau) = \begin{cases} \pi\,, & \tau = -1\,, \\ \displaystyle\int_\tau^1 \frac{dt}{\sqrt{1-t^2}}\,, & \tau \in (-1, 1)\,, \\ 0\,, & \tau = 1\,. \end{cases} \tag{3.9}$$

Das Integral in (3.9) ist ein uneigentliches Riemann-Integral. Die Substitution $\xi = \cos(t)$ liefert die behauptete Integraldarstellung des Arcuscosinus. Siehe Analysis-Skript, Abschnitt 17. Der Arcuscosinus ist auf $[-1, 1]$ streng monoton fallend.

(3) Es gilt

$$\measuredangle(x, y) = \measuredangle\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right)\,.$$

Der unorientierte Winkel zwischen x und y ist gleich dem unorientierten Winkel zwischen den entsprechenden Richtungsvektoren $x/\|x\|$ und $y/\|y\|$.

Im Folgenden begründen wir, warum eine Deutung als Sinus nicht sinnvoll ist. Wir heben ausdrücklich hervor, dass wir Definition und Rechenregeln für Cosinus und Sinus der Analysis entlehnen. Insbesondere übernehmen wir die wichtige Formel

$$\cos^2(\xi) + \sin^2(\xi) = 1\,, \quad \xi \in \mathbb{R}\,. \tag{3.10}$$

Die Formel (3.10) ist eine analytische Version des *Satzes des Pythagoras*. In Satz 3.16 geben wir eine geometrische Version mit Hilfe der Norm.

Der Spezialfall $x = y \neq 0$ zeigt, dass eine alternative Deutung der rechten Seite von Definition (3.7) wegen

$$\cos(0) = 1\,, \quad \sin(0) = 0$$

als Sinus des eingeschlossenen Winkels geometrisch keinen Sinn macht.

Außerdem ist es geometrisch evident, dass beim Ersetzen von x durch $-x$ der eingeschlossene Winkel $\varphi \in [0, \pi]$ durch den Winkel $\pi - \varphi \in [0, \pi]$ ersetzt wird. Die Additionstheoreme

$$\cos(\pi - \varphi) = -\cos(\varphi)\,, \quad \sin(\pi - \varphi) = \sin(\varphi)\,.$$

zeigen, dass eine Deutung der rechten Seite von (3.7) als Sinus im Gegensatz zur Deutung als Cosinus keinen Sinn macht.

Bevor wir den Zusammenhang zwischen (3.7) und der elementargeometrischen Einführung des Cosinus herstellen können, müssen wir den Begriff der *Orthogonalität* einführen. Wir erinnern an

$$\cos(\tfrac{\pi}{2}) = 0\,, \quad \sin(\tfrac{\pi}{2}) = 1\,.$$

Definition 3.13. *Zwei Spaltenvektoren x, $y \in \mathbb{R}^n$ heißen* orthogonal *oder* senkrecht*, wenn*

$$\langle x, y \rangle = x \cdot y = 0 \tag{3.11}$$

gilt. Wir sagen dann auch, dass x zu y orthogonal oder senkrecht ist. Wenn x zu y orthogonal ist, ist y zu x orthogonal. Der Nullvektor $0 \in \mathbb{R}^n$ ist zu allen Vektoren des $\mathbb{R}^n$ orthogonal.

Das folgende Beispiel stellt den Zusammenhang von Definition (3.7) mit der elementargeometrischen Einführung des Cosinus eines Winkels her.

Beispiel 3.14. Wir betrachten in der Ebene $\mathbb{R}^2$ das rechtwinklige Dreieck mit den Eckpunkten 0 und x, y mit

$$x = \begin{pmatrix} \cos(\varphi) \\ 0 \end{pmatrix}\,, \quad y = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}\,, \quad 0 < \varphi < \tfrac{\pi}{2}\,.$$

Die Verbindungsstrecke zwischen 0 und y ist die Hypothenuse, die zwischen 0 und x ist die Ankathete des Dreiecks. Am Eckpunkt x liegt der rechte Winkel des Dreiecks.

$$\langle y - x, 0 - x \rangle = \begin{pmatrix} 0 \\ \sin(\varphi) \end{pmatrix} \cdot \begin{pmatrix} -\cos(\varphi) \\ 0 \end{pmatrix} = 0\,.$$

Die Normen

$$\|y\| = \sqrt{\cos^2(\varphi) + \sin^2(\varphi)} = 1\,, \quad \|x\| = \sqrt{\cos^2(\varphi)} = |\cos(\varphi)| = \cos(\varphi)$$

sind die Längen von Hypothenuse respektive Ankathete. Einsetzen ergibt

$$\cos(x, y) = \frac{\langle x, y \rangle}{\|x\|\,\|y\|} = \frac{\cos(\varphi) \cdot \cos(\varphi) + 0 \cdot \sin(\varphi)}{\cos(\varphi) \cdot 1} = \cos(\varphi) = \frac{\|x\|}{\|y\|}\,.$$

Damit ist der Zusammenhang zur elementargeometrischen Einführung des Cosinus hergestellt. Mit Hilfe der Analysis kann die Länge glatter Kurven definiert werden. Danach ist die positive reelle Zahl φ die Länge des Kreisbogens vom Radius 1 um $0 \in \mathbb{R}^2$, der die beiden Punkte x und y verbindet. $\square$

Aus (3.7) folgt

$$\langle x, y\rangle = \cos(x, y)\, \|x\|\, \|y\| \,. \tag{3.12}$$

Die linke Seite von (3.12) ist nach Satz 3.4 bilinear in den Argumenten x und y. Aus dieser Formel folgen der Cosinussatz 3.15 und als Spezialfall der Satz 3.16 des Pythagoras.

Satz 3.15 (Cosinussatz)**.** *Es seien x, $y \in \mathbb{R}^n$ beliebige vom Nullvektor verschiedene Spaltenvektoren. Dann gelten die beiden Formeln*

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\|\,\cos(x, y)\,,\\ \|x-y\|^2 &= \|x\|^2 + \|y\|^2 - 2\,\|x\|\,\|y\|\,\cos(x, y)\,.\end{aligned}$$

Dafür schreiben wir zusammenfassend

$$\|x \pm y\|^2 = \|x\|^2 + \|y\|^2 \pm 2\,\|x\|\,\|y\|\,\cos(x, y)\,. \tag{3.13}$$

Beweis. Aus (3.3) und (3.12) folgt

$$\begin{aligned}\|x \pm y\|^2 &= \langle x \pm y, x \pm y\rangle\\ &= \langle x, x\rangle \pm 2\,\langle x, y\rangle + \langle y, y\rangle\\ &= \|x\|^2 \pm 2\,\cos(x, y)\,\|x\|\,\|y\| + \|y\|^2\,.\end{aligned}$$

Damit ist der Beweis des Cosinussatzes beendet. □

Satz 3.16 (Satz des Pythagoras)**.** *Es seien x, $y \in \mathbb{R}^n$ beliebige Spaltenvektoren. Dann gilt*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \tag{3.14}$$

genau dann, wenn x und y orthogonal sind.

Beweis. Die Formel (3.14) gilt trivialerweise, wenn einer der beiden Vektoren der Nullvektor ist. In diesem Fall sind x und y orthogonal.

Wir nehmen an, dass x und y vom Nullvektor verschiedene orthogonale Vektoren sind. Dann gilt $\cos(x, y) = 0$. Aus (3.13) folgt damit

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2\,.$$

Wir nehmen nun an, dass x und y vom Nullvektor verschiedene Vektoren sind, die (3.14) erfüllen. Dann gilt

$$0 = \|x+y\|^2 - \|x\|^2 - \|y\|^2 = 2\,\langle x, y\rangle\,.$$

Also sind x und y orthogonal. □

Die Formel $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ gilt genau dann, wenn die Vektoren x und y orthogonal sind. In diesem Fall ist das Paar (x, y) eine orthogonale Zerlegung des Vektors $x + y$. Wir formulieren eine Variante des Satzes von Pythagoras.

Satz 3.17 (Orthogonale Zerlegungen, orthogonale Projektionen)**.** *Seien a, $b \in \mathbb{R}^n$ mit $b \neq 0$ gegeben. Dann sind*

$$v = \frac{\langle a, b\rangle}{\langle b, b\rangle}\, b\,, \quad w = a - \frac{\langle a, b\rangle}{\langle b, b\rangle}\, b \tag{3.15}$$

die einzigen Vektoren v, $w \in \mathbb{R}^n$, die (1), (2), (3) *erfüllen.*

(1) $a = v + w$.

(2) $v \in \mathbb{R}b = \{\lambda b \mid \lambda \in \mathbb{R}\}$.

(3) $\langle w, b\rangle = 0$.

Der Vektor v heißt die orthogonale Projektion von a in Richtung des Vektors b. *Das Paar (v, w) heißt die* orthogonale Zerlegung von a bezüglich b. *Die folgende geometrische Sprechweise ist üblich:* Der Punkt v ist der Fußpunkt des Lotes des Punktes a auf die Gerade $\mathbb{R}b$. Dabei ist die Norm $\|w\|$ die Länge des Lotes.

Beweis. Erster Schritt. Wir setzen (3.15) voraus. Dann sind (1) und (2) offenbar erfüllt. Wegen

$$\begin{aligned}\langle w, b\rangle &= \langle a - \frac{\langle a, b\rangle}{\langle b, b\rangle}\, b, b\rangle \\ &= \langle a, b\rangle - \frac{\langle a, b\rangle}{\langle b, b\rangle}\,\langle b, b\rangle \\ &= \langle a, b\rangle - \langle a, b\rangle = 0\end{aligned}$$

gilt auch (3).

Zweiter Schritt. Wir setzen (1), (2), (3) voraus. Wegen (2) gibt es $\lambda \in \mathbb{R}$ mit $v = \lambda b$. Aus (1) und (3) folgt dann

$$\begin{aligned}\langle a, b\rangle &= \langle v + w, b\rangle \\ &= \langle \lambda b + w, b\rangle \\ &= \langle \lambda b, b\rangle + \langle w, b\rangle \\ &= \lambda\langle b, b\rangle\,.\end{aligned}$$

Wegen $b \neq 0$ gilt $\langle b, b\rangle \neq 0$. Also gelten die Formeln (3.15). Damit ist auch die Einzigkeitsaussage bewiesen. □

Beispiel 3.18 (Orthogonale Zerlegung)**.** Wir berechnen für die beiden Spaltenvektoren $a,\, b \in \mathbb{R}^3$ mit

$$a = \begin{pmatrix} 4 \\ -5 \\ 3 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

die orthogonale Zerlegung von a nach b. Mit Hilfe von (3.15) erhalten wir

$$v = \frac{\langle a, b\rangle}{\langle b, b\rangle}\, b = \frac{8-5-6}{4+1+4} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

und

$$w = a - v = \begin{pmatrix} 4 \\ -5 \\ 3 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 12 \\ -15 \\ 9 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 14 \\ -14 \\ 7 \end{pmatrix}.$$

Nach Konstruktion gilt $a = v + w$. Wir rechnen nach, das die Vektoren v und w zueinander orthogonal sind. Es gilt

$$\langle v, w\rangle = -\frac{28-14-14}{9} = 0\,.$$

Siehe Beispiel 5.7. □

Satz 3.19 (Parallelogrammregel)**.** *Für alle $x,\, y \in \mathbb{R}^n$ gilt*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)\,. \tag{3.16}$$

Beweis. Aus Definition (3.3) folgen

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2\,, \\ \|x-y\|^2 &= \|x\|^2 - 2\langle x, y\rangle + \|y\|^2\,. \end{aligned}$$

Addition beider Gleichungen liefert die Parallelogrammregel (3.16). □

Satz 3.20 (Polarisierung)**.** *Für alle $x,\, y \in \mathbb{R}^n$ gilt*

$$\langle x, y\rangle = \tfrac{1}{4}(\|x+y\|^2 - \|x-y\|^2)\,. \tag{3.17}$$

Beweis. Aus Definition (3.3) folgen

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2\,, \\ \|x-y\|^2 &= \|x\|^2 - 2\langle x, y\rangle + \|y\|^2\,. \end{aligned}$$

Subtraktion beider Gleichungen liefert die Polarisierungsformel (3.17). □

Aus der euklidischen Norm lässt sich ein Abstand zwischen Spaltenvektoren konstruieren.

Definition 3.21. *Der* euklidische Abstand *zweier beliebiger Spaltenvektoren* x, $y \in \mathbb{R}^n$ *wird durch*

$$d(x, y) = \|x - y\| \tag{3.18}$$

definiert. Die Abbildung $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *heißt auch* euklidische Metrik. *Nach Definition (3.3) der euklidischen Norm bedeutet (3.18) ausgeschrieben, dass*

$$d(x, y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}} = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} \tag{3.19}$$

gilt.

An Formel (3.19) können wir ablesen, dass die euklidsische Metrik den Satz des Pythagoras codiert.

Beispiel 3.22. *Für die Spaltenvektoren*

$$x = \begin{pmatrix} 6 \\ 3 \end{pmatrix}, \quad y = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

erhalten wir

$$\begin{aligned} d(x, y) &= \sqrt{(6-7)^2 + (3-5)^2} = \sqrt{1+4} = \sqrt{5}\,, \\ d(x, 0) &= \sqrt{6^2 + 3^2} = \sqrt{36+9} = \sqrt{45} = \|x\|\,, \\ d(y, 0) &= \sqrt{7^2 + 5^2} = \sqrt{49+25} = \sqrt{74} = \|y\|\,. \end{aligned}$$

Die Abschätzungen

$$8.6 < \sqrt{74} < 8.7\,, \quad 6.7 < \sqrt{45} < 6.8\,, \quad 2.2 < \sqrt{5} < 2.4$$

lassen sich leicht durch Quadrieren nachrechnen. Es folgt

$$\sqrt{74} - \sqrt{45} < 8.7 - 6.7 = 2.0 < \sqrt{5}\,.$$

Damit ergibt sich

$$|\,\|x\| - \|y\|\,| = |\sqrt{45} - \sqrt{74}| \le d(x, y) = \sqrt{5} \le \sqrt{45} + \sqrt{74} = \|x\| + \|y\|\,.$$

□

Die Aussagen des folgenden Satzes 3.23 sind einfache Konsequenzen aus der Definition (3.18) der euklidischen Metrik und dem Satz 3.10 über die Eigenschaften der euklidischen Norm.

Satz 3.23. *Seien $x,\ y,\ z \in \mathbb{R}^n$ beliebig gegeben. Dann gelten die folgenden Aussagen.*

(1) $d(x,y) \geq 0$.

(2) $d(x,y) = 0 \Longleftrightarrow x = y$.

(3) $d(x,y) = d(y,x)$.

(4) $d(x,0) = d(0,x) = \|x\|$.

(5) $|\,\|x\| - \|y\|\,| \leq d(x,y) \leq \|x\| + \|y\|$.

(6) $d(x,z) \leq d(x,y) + d(y,z)$.

(7) $d(x+z, y+z) = d(x,y)$.

Dabei ist $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ die euklidische Metrik und $\|\,.\,\| : \mathbb{R}^n \to \mathbb{R}$ die euklidische Norm.

Beweis. Die Aussagen (1) bis (7) folgen aus der Definition (3.18) und den Eigenschaften der euklidischen Norm, die in Satz 3.10 ausgesprochen worden sind. Wir arbeiten die Behauptungen der Reihe ab.

Nachweis von (1). Satz 3.10, Aussage (1) liefert

$$d(x,y) = \|x - y\| \geq 0\,.$$

Nachweis von (2). Satz 3.10, Aussage (2) liefert

$$\begin{aligned} d(x,y) = 0 \quad &\Longleftrightarrow \quad \|x-y\| = 0 \\ &\Longleftrightarrow \quad x - y = 0 \quad \Longleftrightarrow \quad x = y\,. \end{aligned}$$

Nachweis von der Symmetrie (3). Satz 3.10, Aussage (4) liefert

$$d(x,y) = \|x-y\| = \|-(x-y)\| = \|y-x\| = d(y,x)\,.$$

Nachweis von (4). Aus (3) und Satz 3.10, Aussage (4) folgt

$$d(x,0) = d(0,x) = \|0 - x\| = \|-x\| = \|x\|\,.$$

Nachweis von (5). Die Aussagen (6), (4), (5) des Satzes 3.10 liefern

$$|\,\|x\| - \|y\|\,| \leq d(x,y) \equiv \|x-y\| \leq \|x\| + \|y\|\,.$$

Nachweis der Dreiecks-Ungleichung (6). Aus der Dreiecks-Ungleichung für die euklidische Norm nach Aussage (5) des Satzes 3.10 folgt

$$\begin{aligned} d(x,z) = \|x-z\| &= \|(x-y) + (y-z)\| \\ &\leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z)\,. \end{aligned}$$

Nachweis der Translationsinvarianz (7) der euklidischen Metrik.

$$d(x+z, y+z) = \|(x+z) - (y+z)\| = \|x-y\| = d(x,y)\,.$$

□

4 Ergänzung: Hesse-Normalform

Als Anwendung der inneren Produktes behandeln wir die Hesse-Normalform von Hyperebenen. Wir beginnen mit der Erörterung eines Beispieles in der Ebene.

- Gegeben sei die Gerade
$$g_1 \subseteq \mathbb{R}^2 : \quad 6x_1 + 7x_2 = 18 \,.$$
- Gesucht sind
  (1) der euklidische Abstand $d(0, g_1)$ des Nullvektors $0 \in \mathbb{R}^2$ und
  (2) der euklidische Abstand $d(P_1, g_1)$ des Punktes $P_1 = (-2, -3)^t \in \mathbb{R}^2$

  von der Geraden g_1.

Zur Lösung dieser Probleme führen wir die Hesse-Normalform ein. Wir ersetzen die gegebene Geradengleichung

$$\begin{pmatrix} 6 \\ 7 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 6x_1 + 7x_2 = 18 \tag{4.1}$$

durch eine äquivalente Gleichung der Form

$$\langle n, x \rangle = \delta \,, \quad \|n\| = 1 \,, \quad \delta \geq 0 \,. \tag{4.2}$$

Entsprechend verfahren wir mit der Gleichung

$$g \subseteq \mathbb{R}^2 : \quad \alpha_1 x_1 + \alpha_2 x_2 = \gamma \,, \quad \alpha_1^2 + \alpha_2^2 \neq 0$$

einer beliebigen Geraden $g \subseteq \mathbb{R}^2$. In den obigen Gleichungen repräsentieren die Spaltenvektoren $x = (x_1, x_2)^t$ die Punkte, die auf der betrachteten Geraden liegen. Eine Geradengleichung der Form (4.2) heißt *Hesse-Normalform*.

Offenbar gilt $\delta = 0$ genau dann, wenn die Gerade $g \subseteq \mathbb{R}^2$ durch den Nullvektor $0 \in \mathbb{R}^2$ läuft. In diesem Fall ist der Einheitsvektor $n \in \mathbb{R}^2$ bis auf das Vorzeichen eindeutig bestimmt. Im Fall $\delta \neq 0$ ist n dagegen eindeutig bestimmt.

Die beiden Bestimmungsstücke n und δ der Hesse-Normalform lassen sich geometrisch deuten:

$$n = \textit{Normalenrichtung zu } g\textit{, die vom Nullvektor wegweist}\,, \tag{4.3}$$
$$\delta = \textit{euklidischer Abstand zwischen dem Nullvektor und } g\,. \tag{4.4}$$

Folglich legen n und δ die Lage der Geraden g in der Ebene $\mathbb{R}^2$ fest. Später ersetzen wir (4.4) durch die allgemeinere Abstandsformel (4.7)

Für die Geradengleichung der Gerade g_1 erhalten wir aus

$$n = \frac{1}{\sqrt{36+49}} \begin{pmatrix} 6 \\ 7 \end{pmatrix} = \frac{1}{\sqrt{85}} \begin{pmatrix} 6 \\ 7 \end{pmatrix} \,, \quad \delta = \frac{18}{\sqrt{85}} = d(0, g_1) \,.$$

die Hesse-Normalform

$$\tfrac{6}{\sqrt{85}}\,x_1 + \tfrac{7}{\sqrt{85}}\,x_2 = \tfrac{18}{\sqrt{85}}\,. \tag{4.5}$$

Es gilt die Einschließung

$$1.952 < \tfrac{18}{\sqrt{85}} < 1.953\,.$$

Damit ist Problem (1) gelöst. Wir heben ausdrücklich hervor, dass die beiden Gleichungen (4.1) und (4.5) äquivalent sind. Dagegen lässt sich die geometrische Lage der Geraden g_1 in Ebene $\mathbb{R}^2$ ohne weitere Rechnung allein an den Koeffizienten der Hesse-Normalform (4.5) ablesen.

Wir wenden uns nun der Begründung der in (4.3) und (4.4) ausgesprochenen geometrischen Deutung zu. Offenbar gilt

$$\langle \delta n, n\rangle = \delta\langle n, n\rangle = \delta\,, \quad \delta n \in g\,.$$

Die Gerade g kann auch durch eine Gleichung der Form

$$g: \quad x = tw + \delta n\,, \quad \|w\| = 1\,, \quad t \in \mathbb{R}$$

dargestellt werden. Dabei ist der Einheitsvektor $w \in \mathbb{R}^2$ ein *Richtungsvektor* und $t \in \mathbb{R}$ ein *reeller Parameter*, der die Punkte auf der Geraden g parametrisiert. Wie wir gesehen haben, ist δn ein Punkt auf der Geraden. Der Punkt δn wird durch $t = 0$ dargestellt. Es ist klar, dass das Paar (w, t) durch das Paar $(-w, -t)$ ersetzt werden kann. Einsetzen der Parameterdarstellung in die Hesse-Normalform ergibt

$$\begin{aligned}
& (\forall t \in \mathbb{R}): \quad \delta = \langle n, x\rangle = \langle n, tw + \delta n\rangle = t\,\langle n, w\rangle + \delta \\
\Longrightarrow\quad & (\forall t \in \mathbb{R}): \quad 0 = t\,\langle n, w\rangle \\
\Longrightarrow\quad & \langle n, w\rangle = 0\,.
\end{aligned}$$

Die Einheitsvektoren n und w sind demnach orthogonal. Folglich schneiden sich die beiden Geraden g und

$$h: \quad x = tn\,, \quad t \in \mathbb{R}$$

im Punkte $\delta n \in \mathbb{R}^2$ orthogonal. Damit ist die Deutung (4.3) gerechtfertigt.

Weil n normiert ist, gilt

$$\delta = \|\delta n\| = d(\delta n, 0) = d(0, \delta n)\,.$$

Weil die normierten Vektoren n und w orthogonal sind, gilt nach dem Satz 3.16 des Pythagoras

$$\|x\|^2 = \|tw\|^2 + \|\delta n\|^2 = t^2 + \delta^2 \geq \delta^2$$

für jeden Parameterwert $t \in \mathbb{R}$. Also ist δn unter allen Punkten, die auf der Geraden g liegen, der eindeutig bestimmte Punkt mit dem kleinsten Abstand zum Nullvektor $0 \in \mathbb{R}^2$. *Der euklidische Abstand zwischen einem beliebigen Punkt $P \in \mathbb{R}^2$ und der Geraden $g \subseteq \mathbb{R}^2$ wird durch die Gleichung*

$$d(P, g) \equiv \min\{\, d(P, x) \mid x \in g \,\} \tag{4.6}$$

definiert. Nach den obigen Überlegungen gilt

$$d(0, g) = \min\{\, d(0, x) \mid x \in g \,\} = d(0, \delta n) = \delta \,.$$

Damit ist auch die Deutung (4.4) gerechtfertigt.

Wir wenden uns dem Problem (2) zu. Sei $P = (p_1, p_2)^t \in \mathbb{R}^2$ irgendein Punkt der Ebene. Für alle $t \in \mathbb{R}$ gilt

$$\begin{aligned} \|P - (tw + \delta n)\|^2 &= \langle P - (tw + \delta n), P - (tw + \delta n)\rangle \\ &= (\|P\|^2 - 2\delta\langle P, n\rangle + \delta^2) + (t^2 - 2\langle P, w\rangle t)\,. \end{aligned}$$

Wir nehmen die Analysis zu Hilfe. Die Summe in der ersten Klammer auf der rechten Seite der Gleichung hängt nicht von dem reellen Parameter t ab. Die quadratische Funktion

$$t \mapsto t^2 - 2\langle P, w\rangle t$$

nimmt bei $t = \langle P, w\rangle$ das Minimum $-\langle P, w\rangle^2$ an. Die Einheitsvektoren n und w sind orthogonal. Also gelten

$$n = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}, \quad n_1^2 + n_2^2 = \|n\|^2 = 1\,, \quad w = \pm \begin{pmatrix} n_2 \\ -n_1 \end{pmatrix}.$$

Folglich besitzt P die orthogonale Zerlegung

$$\begin{aligned} P &= \left\langle \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \begin{pmatrix} n_2 \\ -n_1 \end{pmatrix} \right\rangle \begin{pmatrix} n_2 \\ -n_1 \end{pmatrix} + \left\langle \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \right\rangle \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \\ &= \langle P, w\rangle w + \langle P, n\rangle n\,. \end{aligned}$$

Damit folgt

$$\begin{aligned} (d(P, g))^2 &= \|P - (\langle P, w\rangle w + \delta n)\|^2 \\ &= \|(P - \langle P, w\rangle w) - \delta n\|^2 \\ &= \|\langle P, n\rangle n - \delta n\|^2 \\ &= |\langle P, n\rangle - \delta|^2 \cdot \|n\|^2 = |\langle P, n\rangle - \delta|^2\,. \end{aligned}$$

Der Übergang zu den positiven Quadratwurzeln liefert

$$d(P, g) = |\langle n, P\rangle - \delta|\,. \tag{4.7}$$

Nach der Abstandsformel (4.7) *gehen in die Berechnung des euklidischen Abstandes $d(P, g)$ zwischen P und g aus der Beschreibung der Geraden g nur die Bestimmungsstücke n und δ der Hesse-Normalform*

$$g \subseteq \mathbb{R}^2 : \quad \langle n, x\rangle = \delta\,, \quad \|n\| = 1\,, \quad \delta \geq 0$$

ein. Dieser Sachverhalt unterstreicht noch einmal die geometrische Bedeutung der Hesse-Normalform. Im Fall $P = 0$ liefert (4.7) die frühere Formel (4.4).

Abschließend berechnen wir $d(P_1, g_1)$ für

$$P_1 = \begin{pmatrix} -2 \\ -3 \end{pmatrix}, \quad g_1 : \quad \tfrac{6}{\sqrt{85}}\, x_1 + \tfrac{7}{\sqrt{85}}\, x_2 = \tfrac{18}{\sqrt{85}} \, .$$

Die Abstandsformel (4.7) liefert

$$d(P_1, g_1) = |\langle n, P_1\rangle - \delta| = |(-\tfrac{12}{\sqrt{85}} - \tfrac{21}{\sqrt{85}}) - \tfrac{18}{\sqrt{85}}| = \tfrac{51}{\sqrt{85}} \, .$$

Es gilt die Einschließung

$$5.531 < \tfrac{51}{\sqrt{85}} < 5.532 \, .$$

Damit sind die Ausführungen zum Beispiel beendet.

Die Überlegungen zur Hesse-Normalform lassen sich auf Hyperebenen des $\mathbb{R}^m$ verallgemeinern.

Definition 4.1.

(1) *Eine Teilmenge $H \subseteq \mathbb{R}^m$ heißt* Hyperebene des $\mathbb{R}^m$*, wenn es einen Spaltenvektor $w \in \mathbb{R}^m \setminus \{0\}$ und ein $\gamma \in \mathbb{R}$ derart gibt, dass*

$$H = \{x \in \mathbb{R}^m : \langle w, x\rangle = \gamma\}$$

gilt. Wir schreiben dann auch

$$H \subseteq \mathbb{R}^m : \quad \langle w, x\rangle = \gamma \, , \quad x \in \mathbb{R}^m \, .$$

Wenn keine Verwechslungen zu befürchten sind, schreiben wir einfach

$$\langle w, x\rangle = \gamma \, .$$

(2) *Für jedes $\lambda \in \mathbb{R} \setminus \{0\}$ beschreibt das Paar $(\lambda w, \lambda\gamma)$ dieselbe Hyperebene H wie das Paar (w, γ). Daher kann ohne Einschränkung w als Einheitsvektor und $\gamma \geq 0$ gewählt werden. Siehe Satz* 4.2.

(3) *Eine Darstellung der Form*

$$H \subseteq \mathbb{R}^m : \quad \langle n, x\rangle = \delta \, , \quad x \in \mathbb{R}^m \, ,$$

wobei $n \in \mathbb{R}^m$ ein Spaltenvektor mit $\|n\| = 1$ ist und $\delta \geq 0$ gilt, heißt eine Hesse-Normalform von H*. Wegen der Symmetrie des inneren Produktes schreiben wir $\langle x, n\rangle = \delta$ an Stelle von $\langle n, x\rangle = \delta$.*

Im Fall $\delta \neq 0$ ist die Hesse-Normalform eindeutig durch die Hyperebene H bestimmt. Im Fall $\delta = 0$ wird H durch die beiden Hesse-Normalformen $\langle n, x\rangle = 0$ und $\langle -n, x\rangle = 0$ mit $\|n\| = 1$ beschrieben.

Satz 4.2. *Zwei Gleichungen*

$$\langle w_i, x\rangle = \gamma_i$$

mit $w_i \in \mathbb{R}^m \setminus \{0\}$ und $\gamma_i \in \mathbb{R}$ für $i = 1, 2$ beschreiben genau dann dieselbe Hyperebene $H \subseteq \mathbb{R}^m$, wenn es ein $\lambda \in \mathbb{R} \setminus \{0\}$ mit

$$\lambda w_1 = w_2 \, , \quad \lambda\gamma_1 = \gamma_2$$

gibt.

Beweis. Nach Multiplikation mit geeigneten von Null verschiedenen reellen Zahlen können wir es so einrichten, dass die Bedingungen $\|w_1\| = \|w_2\| = 1$ und $\gamma_1 \geq 0$ und $\gamma_2 \geq 0$ erfüllt sind.

Fall A. Wir nehmen an, dass $0 \notin H$ gilt. Dann gilt

$$\gamma_i > 0\,, \quad 0 \neq \gamma_i w_i \in H$$

für $i = 1, 2$. Folglich sind die Gleichungen

$$\langle \gamma_2 w_2, w_1 \rangle = \gamma_1\,, \quad \langle \gamma_1 w_1, w_2 \rangle = \gamma_2$$

erfüllt. Addition liefert unter Beachtung der Symmetrie des inneren Produktes

$$(\gamma_1 + \gamma_2)\langle w_1, w_2 \rangle = \gamma_1 + \gamma_2 \neq 0\,.$$

Auflösen nach $\langle w_1, w_2 \rangle$ ergibt

$$\langle w_1, w_2 \rangle = 1\,, \quad \cos(w_1, w_2) = \frac{\langle w_1, w_2 \rangle}{\|w_1\|\,\|w_2\|} = 1\,.$$

Aus dem Cosinussatz folgt

$$\|w_1 - w_2\|^2 = \|w_1\|^2 + \|w_2\|^2 - 2\,\|w_1\|\,\|w_2\|\,\cos(w_1, w_2) = 1 + 1 - 2 = 0\,.$$

Aus der positiven Definitheit der euklidischen Norm folgt $w_1 = w_2$. Wegen

$$\langle \gamma_2 w_2, w_1 \rangle = \gamma_1$$

gilt daher $\gamma_1 = \gamma_2$. In Fall A gilt also $\lambda = 1$.

Fall B. Wir nehmen an, dass $0 \in H$ gilt. Wir zeigen, dass

$$H' = w_1 + H = \{y \in \mathbb{R}^m : ((\exists x \in H) : y = w_1 + x)\}$$

eine Hyperebene mit $0 \notin H'$ und den äquivalenten darstellenden Gleichungen

$$\langle y, w_1 \rangle = 1\,, \quad \langle y, w_2 \rangle = \langle w_1, w_2 \rangle\,, \quad y \in \mathbb{R}^m$$

ist. Weil dann außerdem $\langle w_1, w_2 \rangle^2 = 1$ gilt, ergibt sich nach Fall A, dass $w_1 = w_2$ oder $w_1 = -w_2$ gilt.

Erster Schritt: H' wird durch $\langle y, w_1 \rangle = 1$ beschrieben. Sei $y \in H'$. Dann gibt es $x \in H$ mit $y = w_1 + x$. Folglich gilt

$$\langle y, w_1 \rangle = \langle w_1 + x, w_1 \rangle = \langle w_1, w_1 \rangle + \langle x, w_1 \rangle = 1 + 0 = 1\,.$$

Es gelte umgekehrt $\langle y, w_1 \rangle = 1$. Dann gibt es eine orthogonale Zerlegung

$$y = \langle y, w_1 \rangle w_1 + x = w_1 + x\,, \quad \langle x, w_1 \rangle = 0\,.$$

Das Bestehen der Gleichung $\langle x, w_1 \rangle = 0$ bedeutet, dass x in H liegt. Folglich gehört $y = w_1 + x$ zu H'.

Zweiter Schritt: H' wird durch $\langle y, w_2\rangle = \langle w_1, w_2\rangle$ beschrieben. Sei $y \in H'$. Dann gibt es $x \in H$ mit $y = w_1 + x$. Folglich gilt

$$\langle y, w_2\rangle = \langle w_1 + x, w_2\rangle = \langle w_1, w_2\rangle + \langle x, w_2\rangle = \langle w_1, w_2\rangle + 0 = \langle w_1, w_2\rangle\,.$$

Wir zeigen, dass für alle $y \in \mathbb{R}^m$ mit $\langle y, w_2\rangle = \langle w_1, w_2\rangle$ die Gleichung

$$\langle y, w_1\rangle = \langle w_1, w_2\rangle^2$$

erfüllt ist. Speziell für $y = w_1$ ergibt sich

$$\langle w_1, w_2\rangle^2 = 1\,.$$

Nach dem ersten Schritt liegen alle $y \in \mathbb{R}^n$ mit $\langle y, w_2\rangle = \langle w_1, w_2\rangle$ in der Hyperebene H'.

Sei $y \in \mathbb{R}^m$ mit $\langle y, w_2\rangle = \langle w_1, w_2\rangle$ beliebig gegeben. Es gibt eine orthogonale Zerlegung

$$y = \langle y, w_2\rangle w_2 + x = \langle w_1, w_2\rangle w_2 + x\,, \quad \langle x, w_2\rangle = 0\,.$$

Wegen $\langle x, w_2\rangle = 0$ folgt $x \in H$ und damit $\langle x, w_1\rangle = 0$. Innere Multiplikation mit w_1 ergibt

$$\langle y, w_1\rangle = \langle w_1, w_2\rangle\langle w_2, w_1\rangle + \langle x, w_1\rangle = \langle w_1, w_2\rangle^2\,.$$

Speziell für $y = w_1$ ergibt sich

$$\langle w_1, w_2\rangle^2 = 1\,.$$

Damit erhalten wir

$$\langle y, w_1\rangle = \langle w_1, w_2\rangle^2 = 1\,.$$

Nach dem ersten Schritt liegt y in H'.

Dritter Schritt. Die Gleichung $\langle y, w_1\rangle = 1$ ist eine Hesse-Normalform der Hyperebene H'. Sei

$$\epsilon = \operatorname{sgn}\langle w_1, w_2\rangle\,.$$

Dann ist $\langle y, \epsilon w_2\rangle = 1$ ebenfalls eine Hesse-Normalform von H'. Nach Konstruktion gilt $0 \notin H'$. Nach Schritt A gilt $w_1 = \epsilon w_2$. Im Fall B gilt demnach $\lambda = \pm 1$. Der Beweis des Satzes ist damit beendet. □

Satz und Definition 4.3. *Sei $H \subseteq \mathbb{R}^m$ eine Hyperebene und*

$$\langle n, x\rangle = \delta$$

eine Hesse-Normalform von H.

(1) *Sei $\xi \in H$ beliebig gewählt. Dann besteht H aus allen $z \in \mathbb{R}^m$ der Form*

$$z = \xi + y\,, \quad \langle n, y\rangle = 0\,, \quad y \in \mathbb{R}^m\,.$$

(2) *Für alle* ξ_1, $\xi_2 \in H$ *sind die Vektoren* $\xi_1 - \xi_2$ *und* n *orthogonal.*

(3) *Ein Vektor* $y \in \mathbb{R}^m$ *heißt* orthogonal *zu* H, *wenn* y *und* n *linear abhängig sind.*

(4) *Sei* $y \in \mathbb{R}^m$ *gegeben. Dann enthält die Menge*

$$\{\lambda \in \mathbb{R} \mid (\exists x \in H) : d(y, x) = \lambda\}$$

ein minimales Element $d(y, H)$. *Offenbar gilt* $d(y, H) \geq 0$.

(5) *Sei* $y \in \mathbb{R}^m$ *gegeben. Dann heißt das Minimum*

$$d(y, H) = \min\{\lambda \in \mathbb{R} \mid (\exists x \in H) : d(y, x) = \lambda\}$$

der euklidische Abstand *von* y *und* H.

(6) *Es gilt die Abstandsformel*

$$d(y, H) = |\langle n, y\rangle - \delta| \, .$$

Insbesondere für $y = 0$ *ergibt sich*

$$d(0, H) = \delta \, .$$

(7) *Für* $y \in \mathbb{R}^m$ *gilt* $y \in H$ *genau dann, wenn* $d(y, H) = 0$ *erfüllt ist.*

(8) *Sei* $y \in \mathbb{R}^m$ *beliebig gegeben. Dann gilt*

$$y_H = y - (\langle n, y\rangle - \delta)n \in H \, .$$

Der Vektor y_H *ist der einzige Vektor* $\xi \in H$ *mit*

$$d(y, \xi) = d(y, H) \, .$$

(9) *Für* $y \in \mathbb{R}^m$ *gilt* $y_H = y$ *genau dann, wenn* $y \in H$ *gilt.*

(10) *Sei* $y \in \mathbb{R}^m$ *gegeben. Dann heißt der Vektor* $y_H \in H$ *der* Fußpunkt des Lotes *von* y *auf* H.

Satz 4.4. *Seien* $a_1, \ldots, a_k \in \mathbb{R}^m$ *mit* $k \geq m$ *beliebig gegeben. Es gibt genau dann eine eindeutig bestimmte Hyperebene* $H \subseteq \mathbb{R}^m$ *mit* $a_\kappa \in H$ *für* $\kappa = 1, \ldots, k$, *wenn eine der beiden folgenden Alternativen erfüllt ist.*

(1) *Es gibt einen Einheitsvektor* $n \in \mathbb{R}^m$ *derart, dass* $w = n$ *und* $w = -n$ *die einzigen normierten Lösungen von*

$$\langle w, a_\kappa\rangle = 0 \, , \quad \kappa = 1, \ldots, k \, .$$

sind. In diesem Fall gilt $0 \in H$ *und die beiden äquivalenten Gleichungen*

$$\langle n, x\rangle = 0 \, , \quad \langle -n, x\rangle = 0 \, , \quad x \in \mathbb{R}^m$$

sind die Hesse-Normalformen von H.

(2) *Es gibt einen eindeutig bestimmten Spaltenvektor* $w \in \mathbb{R}^m \setminus \{0\}$ *mit*

$$\langle w, a_\kappa \rangle = 1\,, \quad \kappa = 1, \ldots, k\,.$$

In diesem Fall gilt $0 \notin H$ *und*

$$\left\langle \frac{w}{\|w\|}, x \right\rangle = \frac{1}{\|w\|}\,, \quad x \in \mathbb{R}^m$$

ist die einzige Hesse-Normalform von H.

Definition 4.5. *Sei* $H \subseteq \mathbb{R}^m$ *eine Hyperebene. Sei* $\langle x, n \rangle = \delta$ *eine Hesse-Normalform von* H. *Die Teilmengen*

$$H^+_{n,\delta} = \{x \in \mathbb{R}^m : \langle n, x \rangle > \delta\}\,, \quad H^-_{n,\delta} = \{x \in \mathbb{R}^m : \langle n, x \rangle < \delta\}$$

von $\mathbb{R}^m \setminus H$ *heißen die* zu H assozierten offenen Halbräume. *Es gilt*

$$H^+_{n,\delta} \cup H^-_{n,\delta} = \mathbb{R}^m \setminus H\,, \quad H^+_{n,\delta} \cap H^-_{n,\delta} = \emptyset\,.$$

Wir sagen, dass $y, z \in \mathbb{R}^m \setminus H$ *durch* H getrennt *werden, wenn* y *und* z *in verschiedenen zu* H *assozierten offenen Halbräumen enthalten sind.*

Beispiel 4.6. Gegeben seien im $\mathbb{R}^4$ die vier Punkte

$$A = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 4 \\ 3 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ -2 \\ -3 \\ 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 \\ 3 \\ -2 \\ -3 \end{pmatrix}$$

sowie die beiden Punkte

$$P = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

(1) Wir zeigen zuerst, dass es genau eine Hyperebene $H \subseteq \mathbb{R}^4$ gibt, die die vier Punkte A, B, C und D enthält.

Nach Satz 4.4 ist das lineare Gleichungssystem

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 4 & 3 & 1 \\ 1 & -2 & -3 & 0 \\ 0 & 3 & -2 & -3 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} \gamma \\ \gamma \\ \gamma \\ \gamma \end{pmatrix}$$

zu untersuchen. Wir wenden das Verfahren von Gauß-Jordan an.

$$
\begin{array}{rrrr|r|l}
1 & 0 & -1 & 0 & \gamma & \\
2 & 4 & 3 & 1 & \gamma & \\
1 & -2 & -3 & 0 & \gamma & \\
0 & 3 & -2 & -3 & \gamma & \\
\hline
1 & 0 & -1 & 0 & \gamma & Z_1 \\
0 & 4 & 5 & 1 & -\gamma & Z_2 - 2Z_1 \\
0 & -2 & -2 & 0 & 0 & Z_3 - Z_1 \\
0 & 3 & -2 & -3 & \gamma & Z_4 \\
\hline
1 & 0 & -1 & 0 & \gamma & Z_1 \\
0 & 1 & \frac{5}{4} & \frac{1}{4} & -\frac{1}{4}\gamma & \frac{1}{4}Z_2 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2}\gamma & Z_3 + \frac{1}{2}Z_2 \\
0 & 0 & -\frac{23}{4} & -\frac{15}{4} & \frac{7}{4}\gamma & Z_4 - \frac{3}{4}Z_2 \\
\hline
1 & 0 & 0 & 1 & 0 & Z_1 + 2Z_3 \\
0 & 1 & 0 & -1 & \gamma & Z_2 - \frac{5}{2}Z_3 \\
0 & 0 & 1 & 1 & -\gamma & 2Z_3 \\
0 & 0 & 0 & 2 & -4\gamma & Z_4 + \frac{23}{2}Z_3 \\
\hline
1 & 0 & 0 & 0 & 2\gamma & Z_1 - \frac{1}{2}Z_4 \\
0 & 1 & 0 & 0 & -\gamma & Z_2 + \frac{1}{2}Z_4 \\
0 & 0 & 1 & 0 & \gamma & Z_3 - \frac{1}{2}Z_4 \\
0 & 0 & 0 & 1 & -2\gamma & \frac{1}{2}Z_4
\end{array}
$$

Wir erhalten das äquivalente Gleichungssystem

$$
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}
=
\begin{pmatrix} 2\gamma \\ -\gamma \\ \gamma \\ -2\gamma \end{pmatrix} . \tag{4.8}
$$

Es gelten die folgenden beiden Aussagen (i) und (ii).

(i) Für $\gamma = 0$ besitzt (4.8) nur die triviale Lösung $w = 0 \in \mathbb{R}^4$.

(ii) Für $\gamma = 1$ besitzt (4.8) nur die nicht-triviale Lösung

$$
w = \begin{pmatrix} 2 \\ -1 \\ 1 \\ -2 \end{pmatrix} .
$$

Also ist die Alternative (2) aus Satz 4.4 erfüllt. Daher gibt es eine einzige Hyperebene $H \subseteq \mathbb{R}^4$ mit A, B, C, $D \in H$. Es gilt $0 \notin H$.

(2) Das Gleichungssystem (4.8) besitzt für $\gamma = 1$ die eindeutig bestimmte nicht-triviale Lösung $w = (2, -1, 1, -2)^t$. Daher wird H durch die Gleichung

$$2x_1 - x_2 + x_3 - 2x_4 = 1$$

beschrieben. Normieren ergibt die Hesse-Normalform

$$\frac{2}{\sqrt{10}}\,x_1 - \frac{1}{\sqrt{10}}\,x_2 + \frac{1}{\sqrt{10}}\,x_3 - \frac{2}{\sqrt{10}}\,x_4 = \frac{1}{\sqrt{10}}$$

für die Hyperebene H. Wir setzen

$$n = \frac{w}{\|w\|} = \frac{1}{\sqrt{10}} \begin{pmatrix} 2 \\ -1 \\ 1 \\ -2 \end{pmatrix}, \quad \delta = \frac{1}{\sqrt{10}}\,.$$

(3) Wir berechnen nach Satz 4.3 die Abstände der Punkte 0 und P von H.

$$\begin{aligned} d(0, H) &= \delta = \frac{1}{\sqrt{10}}\,, \\ d(P, H) &= |\langle P, n\rangle - \delta| = \left| -\frac{1}{\sqrt{10}} - \frac{1}{\sqrt{10}} \right| = \frac{2}{\sqrt{10}}\,. \end{aligned}$$

(4) Sei $g \subseteq \mathbb{R}^4$ die Gerade durch den Punkt P, die die Hyperebene H orthogonal schneidet. Wir berechnen den Schnittpunkt $P_H \in \mathbb{R}^4$ von g und H.

Die Gerade g besteht aus allen $x \in \mathbb{R}^4$ mit $x = P + tn$, wobei $t \in \mathbb{R}$ ein Parameter ist. Einsetzen in die Hesse-Normalform $\langle x, n\rangle = \delta$ und Auflösen nach t ergibt $t = \delta - \langle P, n\rangle$. Daher folgt

$$P_H = P + (\delta - \langle P, n\rangle)\, n = \frac{1}{5} \begin{pmatrix} 7 \\ 9 \\ 6 \\ 3 \end{pmatrix}.$$

(5) Zum Abschluss untersuchen wir, welche der drei Punkte 0, P, $Q \in \mathbb{R}^4$ durch die Hyperebene H voneinander getrennt werden.

Es gelten

$$\langle 0, n\rangle = 0 < \delta\,, \quad \langle P, n\rangle = -\frac{1}{\sqrt{10}} < \delta\,, \quad \langle Q, n\rangle = \frac{2}{\sqrt{10}} > \delta\,.$$

Also liegen 0 und P im Halbraum $\langle x, n\rangle < \delta$. Dagegen liegt Q im Halbraum $\langle x, n\rangle > \delta$. Daher trennt H den Punkt Q von den Punkten 0 und P. Die beiden Punkte 0 und P werden durch H nicht getrennt.

□

5 Matrizen. Die allgemeine lineare Gruppe

In diesem Abschnitt definieren wir die Multiplikation eines Skalars mit einer Matrix sowie die Addition und die Multiplikation zweier Matrizen. Die genannten Rechenoperationen werden auf die algebraischen Rechenoperationen reeller Zahlen zurückgeführt. Außerdem definieren wir die Transposition von Matrizen. Beim Rechnen mit Matrizen treten neue Phänomene auf.

(1) Die Matrixmultiplikation erfüllt das Kommutativgesetz nicht: Es gibt Matrizen A und B mit $AB \neq BA$.

(2) Es gibt Nullteiler: Es gibt Matrizen $A \neq 0$ und $B \neq 0$ mit $AB = 0$.

(3) Es gibt nilpotente Matrizen: Es gibt Matrizen $A \neq 0$ und $n \in \mathbb{N}$ mit $A^n = 0$.

Diese Phänomene treten bereits beim Rechnen mit (2×2)-Matrizen auf (vergleiche Beispiel 5.15) und müssen bei der Konstruktion von Normalformen für Matrizen berücksichtigt werden. Wie bereits mehrfach gesagt, sind die Rechenoperationen für Matrizen und Spaltenvektoren derart definiert worden, dass sie das Rechnen mit linearen Gleichungssystemen wiedergeben und überschaubar gestalten. Insbesondere lassen sich die elementaren Spalten- und Zeilenumformungen durch die Multiplikation mit Matrizen beschreiben.

Mit $\mathbb{R}^{n\times m}$ bezeichnen wir die Menge der reellen $(n \times m)$-Matrizen. Jede Matrix $A \in \mathbb{R}^{n\times m}$ besteht aus nm *Einträgen* $a_{\nu\mu} \in \mathbb{R}$ mit $\nu = 1, \ldots, n$ und $\mu = 1, \ldots, m$. Die Einträge $a_{\nu\mu}$ werden in einem Rechtecksschema aus n Zeilen und m Spalten angeordnet, wobei der erste Index ν der *Zeilenindex* und der zweite Index μ der *Spaltenindex* ist. Der Lesbarkeit halber wird das Rechtecksschema der Einträge $a_{\nu\mu}$ in Klammern eingeschlossen.

$$A = (a_{\nu\mu}) = (a_{\nu\mu})_{\substack{\nu=1,\ldots,n\\ \mu=1,\ldots,m}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}.$$

Der Eintrag $a_{\nu\mu}$ heißt auch *Eintrag von A an der Position $\nu\mu$*. Manchmal verwenden wir im Zusammenhang mit Matrizen an Stelle des Ausdruckes *Eintrag* den Ausdruck *Matrixeintrag*.

Die Matrix A wird auch als *Zusammenstellung* ihrer m Spaltenvektoren s_μ respektive ihrer n Zeilenvektoren z_ν mit

$$s_\mu = \begin{pmatrix} a_{1\mu} \\ \vdots \\ a_{n\mu} \end{pmatrix} \in \mathbb{R}^n, \quad \mu = 1, \ldots m,$$

$$z_\nu = (a_{\nu 1}, \ldots, a_{\nu m}) \in (\mathbb{R}^m)^t, \quad z_\nu^t = \begin{pmatrix} a_{\nu 1} \\ \ldots \\ a_{\nu m} \end{pmatrix} \in \mathbb{R}^m, \quad \nu = 1, \ldots, n$$

aufgefasst. Dementsprechend schreiben wir

$$A = (s_1, \dots, s_m) = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{R}^{n\times m}\,. \tag{5.1}$$

Wir identifizieren n-komponentigen Spaltenvektoren mit $(n \times 1)$-Matrizen und m-komponentigen Zeilenvektoren mit $(1\times m)$-Matrizen. Schließlich identifizieren wir die reellen (1×1)-Matrizen mit den reellen Zahlen. Um die Identifikation strukturierter Mengen von der Gleichheit zu unterscheiden, verwenden wir das Zeichen $\cong$ und schreiben

$$\mathbb{R}^{n\times 1} \cong \mathbb{R}^n\,, \quad \mathbb{R}^{1\times m} \cong (\mathbb{R}^m)^t \cong \mathbb{R}^m\,, \quad \mathbb{R}^{1\times 1} \cong \mathbb{R}^1 \cong (\mathbb{R}^1)^t \cong \mathbb{R}\,. \tag{5.2}$$

Die in (5.2) auftretenden Mengen tragen eine Vektorraumstruktur. Wir werden später sehen, dass die durch die Beziehungen (5.2) miteinander identifizierten strukturierten Mengen als reelle Vektorräume isomorph sind. *Isomorphe Vektorräume werden als strukturgleiche mathematische Objekte identifiziert.* Die Isomorphie ist ein abgeschwächter Gleichheitsbegriff.

Zwei Matrizen $A = (a_{\nu\mu})$ und $B = (b_{\nu\mu})$ aus der Menge $\mathbb{R}^{n\times m}$ sind nach Definition genau dann *gleich*, wenn $a_{\nu\mu} = b_{\nu\mu}$ für alle $\nu = 1, \dots, n$ und alle $\mu = 1, \dots, m$ gilt. In Zeichen:

$$\begin{aligned} &(\forall A, B \in \mathbb{R}^{n\times m}): \\ &A = B \iff (\forall \nu = 1, \dots n)(\forall \mu = 1, \dots m): \quad a_{\nu\mu} = b_{\nu\mu}\,. \end{aligned} \tag{5.3}$$

Eine $(n \times m)$-Matrix, deren sämtliche Einträge verschwinden, heißt *Nullmatrix* und wird mit $0_{n\times m}$ oder 0 bezeichnet. Die Mengen $\mathbb{R}^{n\times m}$ sind nicht leer, denn sie enthalten die entsprechende Nullmatrix:

$$(\forall n, m \in \mathbb{N}): \quad 0 \equiv 0_{n\times m} \in \mathbb{R}^{n\times m}\,. \tag{5.4}$$

Eine $(n \times m)$-Matrix A mit $n = m$ heißt *quadratische* Matrix. Ein wichtiges Beispiel ist die Einheitsmatrix E. Zur ihrer bequemen Beschreibung dient das Kronecker-Symbol. Für $\nu, \mu = 1, \dots, n$ wird das Kronecker-Symbol $\delta_{\nu\mu}$ durch

$$\delta_{\nu\mu} = \begin{cases} 1\,, & \nu = \mu\,, \\ 0\,, & \nu \neq \mu \end{cases} \tag{5.5}$$

definiert. Die *Einheitsmatrix* $E \in \mathbb{R}^{n\times n}$ besitzt auf der Hauptdiagonalen lauter Einträge gleich 1, während alle anderen Einträge verschwinden. Mit dem Kronecker-Symbol lässt sich die Definition der Einheitsmatrix in der Form

$$E = (\delta_{\nu\mu})_{\substack{\nu=1,\dots,n \\ \mu=1,\dots,n}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{n\times n} \tag{5.6}$$

ausdrücken. Die Spalten von $E \in \mathbb{R}^{n\times n}$ sind die *kanonischen Basisvektoren*

$$e_\mu = (\delta_{\nu\mu}) = (\delta_{\nu\mu})_{\nu=1,\dots,n} \in \mathbb{R}^n\,, \quad \mu = 1,\dots,n\,. \tag{5.7}$$

Ausgeschrieben bedeutet dies

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\,, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}\,, \quad \dots \quad , \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}\,. \tag{5.8}$$

Damit erhalten wir

$$E = (e_1, e_2, \dots, e_n) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}\,. \tag{5.9}$$

Um Verwechslungen mit den kanonischen Basisvektoren anderer Spaltenräume zu vermeiden, schreiben wir für die kanonischen Basisvektoren des $\mathbb{R}^n$ etwas umständlich

$$e_1^{(n)}, \dots, e_n^{(n)}\,.$$

Entsprechend schreiben wir

$$E_n = (e_1^{(n)}, \dots, e_n^{(n)})$$

für die Einheitsmatrix im Matrixraum $\mathbb{R}^{n\times n}$ und

$$\delta_{\nu\mu}^{(n)} = \begin{cases} 1\,, & \nu = \mu\,, \\ 0\,, & \nu \neq \mu \end{cases}$$

für das Kronecker-Symbol, das die Einträge von E_n beschreibt. Vergleiche etwa Satz 5.8 über die Matrixeinträge. Der folgende Satz 5.1 ergibt sich direkt aus den Definitionen.

Satz 5.1. *Seien $e_1, \dots, e_n \in \mathbb{R}^n$ die kanonischen Basisvektoren. Sei $\langle .\,,.\rangle$ das euklidische innere Produkt und $\|\,.\,\|$ die euklidische Norm des $\mathbb{R}^n$. Dann gelten die folgenden Aussagen.*

(1) *Für alle $i, j = 1, \dots, n$ gilt $\langle e_i, e_j\rangle = \delta_{ij}$.*

(2) *Für alle $i = 1, \dots, n$ gilt $\|e_i\| = 1$.*

(3) *Die Menge $\{e_1, \dots, e_n\}$ ist linear unabhängig.*

(4) *Für jeden Spaltenvektor $x = (x_i)_{i=1,\dots,n} \in \mathbb{R}^n$ gilt*

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^{n} x_i e_i\,.$$

(5) *Jeder Spaltenvektor $x \in \mathbb{R}^n$ besitzt eine eindeutig bestimmte Darstellung der Form*

$$x = \sum_{i=1}^{n} \alpha_i e_i \,, \quad \alpha_i \in \mathbb{R} \,.$$

Dabei sind die eindeutig bestimmten Entwicklungskoeffizienten α_i durch

$$\alpha_i = \langle x, e_i \rangle \,, \quad i = 1, \ldots, n$$

gegeben. Wenn $x_1, \ldots, x_n$ die Komponenten des Spaltenvektors x sind, dann gilt $x_i = \alpha_i$ für jeden Index $i = 1, \ldots, n$.

Vermittels der Transposition werden Spalten- in Zeilenvektoren und umgekehrt Zeilen- in Spaltenvektoren überführt. Wenn wir diese Operationen auf die Spalten respektive Zeilen einer $(n \times m)$-Matrix A anwenden, ergibt sich eine $(m \times n)$-Matrix, die mit A^t bezeichnet wird. Die Transposition ist die Abbildung $^t : \mathbb{R}^{n\times m} \to \mathbb{R}^{m\times n}$ mit

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}^t = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \ldots & a_{nm} \end{pmatrix} . \tag{5.10}$$

Die Transposition macht aus dem ν-ten Zeilenindex von A den ν-ten Spaltenindex von A^t und aus dem μ-ten Spaltenindex von A^t den μ-ten Zeilenindex von A^t. In Zeichen:

$$\begin{aligned} &^t : \mathbb{R}^{n\times m} \to \mathbb{R}^{m\times n} \,, \\ &A = (a_{\nu\mu})_{\substack{\nu=1,\ldots,n \\ \mu=1,\ldots,m}} \mapsto A^t = (\tilde{a}_{\mu\nu})_{\substack{\mu=1,\ldots,m \\ \nu=1,\ldots,n}} \,, \quad \tilde{a}_{\mu\nu} = a_{\nu\mu} \,. \end{aligned} \tag{5.11}$$

Für die Matrixeinträge bedeutet dies mit anderen Worten

$$\begin{aligned} &\text{Eintrag von } A^t \text{ an der Position } \mu\nu \\ &\qquad = \text{ Eintrag von } A \text{ an der Position } \nu\mu \,. \end{aligned} \tag{5.12}$$

Mit den Bezeichnungen aus (5.1) lässt sich die Wirkung der Transposition auch in der Form

$$A = (s_1, \ldots, s_m) = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \mapsto A^t = \begin{pmatrix} s_1^t \\ \vdots \\ s_m^t \end{pmatrix} = (z_1^t, \ldots, z_n^t) \tag{5.13}$$

schreiben. Offenbar gilt

$$A^{tt} \equiv (A^t)^t = A \tag{5.14}$$

für alle Matrizen $A \in \mathbb{R}^{n\times m}$.

Beispiel 5.2.

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad E^t = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = E\,.$$

$$A = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 10 & 3 \\ 0 & 3 & -49 \end{pmatrix} = A^t\,.$$

$$B = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}, \quad B^t = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix} \neq B\,.$$

$$C = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}, \quad C^t = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}.$$

□

Seien nun $\alpha \in \mathbb{R}$ und $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ beliebig gegeben. Dann wird die Matrix $\alpha A = \alpha \cdot A \in \mathbb{R}^{n\times m}$ durch

$$\alpha A = \alpha \cdot A = (\alpha a_{\nu\mu})_{\substack{\nu=1,\dots,n \\ \mu=1,\dots,m}} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1m} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2m} \\ \vdots & \vdots & & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \dots & \alpha a_{nm} \end{pmatrix} \tag{5.15}$$

definiert. Insbesondere ergeben sich für die Skalare $\alpha = 0$ und $\alpha = 1$ die Beziehungen

$$0 \cdot A = 0_{n\times m} \in \mathbb{R}^{n\times m}, \quad 1 \cdot A = A \in \mathbb{R}^{n\times m}\,. \tag{5.16}$$

Beispiel 5.3.

$$\alpha E = \alpha \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & \alpha & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \alpha \end{pmatrix}.$$

$$\alpha \begin{pmatrix} 1 & -2 & 3 \\ 0 & 6 & -8 \end{pmatrix} = \begin{pmatrix} \alpha & -2\alpha & 3\alpha \\ 0 & 6\alpha & -8\alpha \end{pmatrix}.$$

$$(-7) \cdot \begin{pmatrix} 1 & -2 & 3 \\ 0 & 6 & -8 \end{pmatrix} = \begin{pmatrix} -7 & 14 & -21 \\ 0 & -42 & 56 \end{pmatrix} .$$

□

Nach Definition wird eine Matrix $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ von links mit einem Skalar $\alpha \in \mathbb{R}$ multipliziert, indem jeder Eintrag $a_{\nu\mu}$ der Matrix A mit α multipliziert wird:

$$\alpha a_{\nu\mu} = \text{ Eintrag von } \alpha A \text{ an der Position } \nu\mu\,. \tag{5.17}$$

Wir erinnern an die entsprechende Bemerkungen zu Definition (2.1) und heben ausdrücklich hervor, dass ein Produkt $A\alpha$ nicht definiert worden ist.

Die $(n \times m)$-Matrix $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ und der Spaltenvektor $x = (x_\mu) \in \mathbb{R}^m$ seien beliebig gegeben. Dann seien die Komponenten y_ν des Spaltenvektors $y = (y_\nu) \in \mathbb{R}^n$ durch die Gleichungen

$$y_\nu \equiv \sum_{\mu=1}^{m} a_{\nu\mu} x_\mu\,, \quad \nu = 1, \ldots, n \tag{5.18}$$

definiert. Wir setzen

$$Ax \equiv y \in \mathbb{R}^n\,. \tag{5.19}$$

Der Spaltenvektor Ax heißt *Bild des Spaltenvektors x unter der Anwendung der Matrix A* oder einfach *Bild von x unter A*. Außerdem bezeichnen wir Ax als Produkt der Matrix A mit dem Spaltenvektor x.

Beispiele 5.4.

$$\begin{pmatrix} 1 & 0 & -3 \\ -1 & 4 & 8 \\ 0 & 1 & 0 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2-0+3 \\ -2-4-8 \\ 0-1-0 \\ 4-2-2 \end{pmatrix} = \begin{pmatrix} 5 \\ -14 \\ -1 \\ 0 \end{pmatrix} .$$

$$\begin{pmatrix} 1 & 0 & -3 \\ -1 & 4 & 8 \\ 0 & 1 & 0 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 8 \\ 0 \\ 2 \end{pmatrix} .$$

$$\begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 4 & 1 & 2 \\ -3 & 8 & 0 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ -3 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4+3+0-2 \\ 0-12+2-2 \\ -12-24+0-2 \end{pmatrix} = \begin{pmatrix} 5 \\ -12 \\ -38 \end{pmatrix} .$$

$$\begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 4 & 1 & 2 \\ -3 & 8 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 8 \end{pmatrix} .$$

$$\begin{pmatrix} 5 \\ -14 \\ -1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ -3 \\ 2 \\ -1 \end{pmatrix} = 20 + 42 - 2 - 0 = 60 \,.$$

$$\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ -12 \\ -38 \end{pmatrix} = 10 + 12 + 38 = 60 \,.$$

□

Matrizen können auf verschiedene Weisen konstruiert werden. Eine wichtige Konstruktion ist das Kronecker-Produkt. Aus einem Spaltenvektor $a \in \mathbb{R}^n$ und einem Spaltenvektor $b \in \mathbb{R}^m$ wird eine $(n \times m)$-Matrix $a \otimes b$ konstruiert. Eine typische Anwendung ist die folgende. Sei $v \in \mathbb{R}^n$ ein Einheitsvektor. Dann wird die orthogonale Projektion des $\mathbb{R}^n$ auf die Gerade $\mathbb{R}v$ durch die Projektionsmatrix $v \otimes v$ beschrieben. Ein erstes Beispiel wird in 5.7 behandelt. Für eine systematische Anwendung des Kronecker-Produktes siehe die Ausführungen zum Spektralsatz in Abschnitt 16. Der Spektralsatz wird in der Quanteninformatik verwendet.

Satz und Definition 5.5 (Kronecker-Produkt)**.** *Seien n, $m \in \mathbb{N}$. Seien $a = (a_\nu)_{\nu=1,\dots,n} \in \mathbb{R}^n$ und $b = (b_\mu)_{\mu=1,\dots,m} \in \mathbb{R}^m$ beliebig gegeben.*

(1) *Die Matrix*

$$a \otimes b = (t_{\nu\mu})_{\substack{\nu=1,\dots,n \\ \mu=1,\dots,m}} \in \mathbb{R}^{n\times m} \,, \quad t_{\nu\mu} = a_\nu b_\mu$$

heißt das Kronecker-Produkt *oder das* Tensorprodukt *von a und b.*

(2) *Für alle $x \in \mathbb{R}^m$ gilt*

$$(a \otimes b)x = \langle x, b\rangle a \in \mathbb{R}^n \,.$$

Dabei ist $\langle \,.\,, .\, \rangle$ das euklidische innere Produkt des $\mathbb{R}^m$.

Beweis. Mit Satz 5.1 erhalten wir

$$(a \otimes b)x = \sum_{\nu=1}^{n} \left(\sum_{\mu=1}^{m} t_{\nu\mu} x_\mu \right) e_\nu^{(n)}$$

$$= \sum_{\nu=1}^{n} \left(\sum_{\mu=1}^{m} a_\nu b_\mu x_\mu \right) e_\nu^{(n)}$$

$$= \left(\sum_{\mu=1}^{m} b_\mu x_\mu \right) \left(\sum_{\nu=1}^{n} a_\nu e_\nu^{(n)} \right)$$

$$= \langle x, b \rangle a \in \mathbb{R}^n .$$

Damit ist Aussage (2) bewiesen. □

Beispiele 5.6.

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 & 1 \cdot 6 & 1 \cdot 7 \\ 2 \cdot 5 & 2 \cdot 6 & 2 \cdot 7 \\ 3 \cdot 5 & 3 \cdot 6 & 3 \cdot 7 \end{pmatrix} = \begin{pmatrix} 5 & 6 & 7 \\ 10 & 12 & 14 \\ 15 & 18 & 21 \end{pmatrix} \in \mathbb{R}^{3 \times 3} .$$

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \cdot 4 & 1 \cdot 5 & 1 \cdot 6 & 1 \cdot 7 \\ 2 \cdot 4 & 2 \cdot 5 & 2 \cdot 6 & 2 \cdot 7 \\ 3 \cdot 4 & 3 \cdot 5 & 3 \cdot 6 & 3 \cdot 7 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 & 7 \\ 8 & 10 & 12 & 14 \\ 12 & 15 & 18 & 21 \end{pmatrix} \in \mathbb{R}^{3 \times 4} .$$

□

Beispiel 5.7 (Projektionsmatrizen)**.** Wir betrachten nocheinmal die beiden Spaltenvektoren

$$a = \begin{pmatrix} 4 \\ -5 \\ 3 \end{pmatrix} , \quad b = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

aus Beispiel 3.18. Die lineare Abbildung, die jedem Spaltenvektor $x \in \mathbb{R}^3$ seinen Fußpunkt

$$\frac{\langle x, b \rangle}{\langle b, b \rangle} b$$

auf der Geraden $\mathbb{R}^3 b$ zuordnet, lässt sich durch eine Matrix $P \in \mathbb{R}^{3 \times 3}$ beschreiben. Mit 5.5 erhalten wir

$$Px = \frac{\langle x, b \rangle}{\langle b, b \rangle} b = \frac{b \otimes b}{\langle b, b \rangle} x = \begin{pmatrix} \frac{b_1 b_1}{\|b\|^2} & \frac{b_1 b_2}{\|b\|^2} & \frac{b_1 b_3}{\|b\|^2} \\ \frac{b_2 b_1}{\|b\|^2} & \frac{b_2 b_2}{\|b\|^2} & \frac{b_2 b_3}{\|b\|^2} \\ \frac{b_3 b_1}{\|b\|^2} & \frac{b_3 b_2}{\|b\|^2} & \frac{b_3 b_3}{\|b\|^2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} .$$

Im vorliegenden Fall gilt

$$P = \frac{1}{9} \cdot \left\{ \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \otimes \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \right\} = \frac{1}{9} \begin{pmatrix} 4 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 4 \end{pmatrix} = \begin{pmatrix} \frac{4}{9} & \frac{2}{9} & -\frac{4}{9} \\ \frac{2}{9} & \frac{1}{9} & -\frac{2}{9} \\ -\frac{4}{9} & -\frac{2}{9} & \frac{4}{9} \end{pmatrix} .$$

Insbesondere gilt

$$Pa = \frac{1}{9} \begin{pmatrix} 4 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 4 \end{pmatrix} \begin{pmatrix} 4 \\ -5 \\ 3 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} .$$

Es gelten $P = P^t$ und $P^2 = P$. Reelle Matrizen mit diesen beiden Eigenschaften heißen *Projektionsmatrizen.* $\square$

Der folgende Satz bringt noch einmal zu Ausdruck, dass Matrizen durch ihre Einträge respektive durch ihre Spalten charakterisiert sind.

Satz 5.8. *Sei $\langle\, .\, ,.\, \rangle$ das euklidische innere Produkt des $\mathbb{R}^n$. Es seien $e_1^{(n)}, \ldots ,$ $e_n^{(n)}$ die kanonischen Basisvektoren von $\mathbb{R}^n$ und $e_1^{(m)}, \ldots , e_m^{(m)}$ die kanonischen Basisvektoren von $\mathbb{R}^m$. Es sei $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ eine beliebige Matrix. Die Spaltenvektoren $a_1, \ldots , a_m \in \mathbb{R}^n$ seien die Spalten der Matrix A. Dann gelten die folgenden Aussagen.*

(1) $(\forall\mu = 1, \ldots , m) : \quad a_\mu = Ae_\mu^{(m)} \in \mathbb{R}^n .$

(2) $(\forall\nu = 1, \ldots , n)(\forall\mu = 1, \ldots , m) : \quad a_{\nu\mu} = \langle e_\nu^{(n)}, Ae_\mu^{(m)} \rangle .$

Beweis. Nachweis von (1). Wir haben für jedes $\mu = 1, \ldots , m$ und jedes $\nu = 1, \ldots , n$ zu zeigen, dass die ν-te Komponente von a_μ mit der ν-ten Komponente von Ae_μ übereinstimmt. Nach Definition gilt

$$a_\mu = (a_{\nu\mu})_{\nu=1,\ldots,n} , \quad e_\mu = (\delta_{\kappa\mu})_{\kappa=1,\ldots,m}$$

für $\mu = 1, \ldots , m$. Aus (5.18) und der Definition (5.5) des Kronecker-Symbols folgt

$$\begin{aligned} &\nu\text{-te Komponente von } Ae_\mu \\ &\quad = \sum_{\kappa=1}^{m} a_{\nu\kappa}\delta_{\kappa\mu} = a_{\nu\mu} = \nu\text{-te Komponente von } a_\mu . \end{aligned}$$

Damit ist Aussage (1) bewiesen.

Nachweis von (2). Aus Aussage (5) des Satzes 5.1 und der soeben bewiesenen Aussage (1) folgt

$$\begin{aligned} a_{\nu\mu} &= \nu\text{-te Komponente von } a_\mu \\ &= \langle a_\mu, e_\nu^{(n)} \rangle = \langle e_\nu^{(n)}, a_\mu \rangle = \langle e_\nu^{(n)}, Ae_\mu^{(m)} \rangle . \end{aligned}$$

Damit ist auch Aussage (2) bewiesen. Der Beweis ist damit beendet. $\square$

Reelle $(n \times m)$-Matrizen definieren lineare Abbildungen des $\mathbb{R}^m$ in den $\mathbb{R}^n$. Später werden wir sehen, dass alle lineare Abbildungen zwischen Vektorräumen endlicher Dimension durch Matrizen beschrieben werden können.

Satz 5.9. *Sei $A \in \mathbb{R}^{n\times m}$ beliebig gegeben. Dann definiert*

$$\varphi(x) = Ax\,, \quad x \in \mathbb{R}^m \tag{5.20}$$

eine lineare Abbildung $\varphi : \mathbb{R}^m \to \mathbb{R}^n$. Die Spalten a_μ und die Matrixeinträge $a_{\nu\mu}$ der Matrix A lassen sich aus der Abbildung φ rekonstruieren.

(1) $(\forall \mu = 1, \ldots, m): \quad a_\mu = \varphi(e_\mu^{(m)}) \in \mathbb{R}^n$.

(2) $(\forall \nu = 1, \ldots, n)(\forall \mu = 1, \ldots, m): \quad a_{\nu\mu} = \langle e_\nu^{(n)}, \varphi(e_\mu^{(m)})\rangle$.

Die Aussagen (1) *und* (2) *sind lediglich Umformulierungen der entsprechenden Aussagen des Satzes* 5.8.

Beweis. Zum Nachweis der Linearität von φ haben wir zu zeigen, dass

$$(\forall \alpha, \beta \in \mathbb{R})(\forall x, y \in \mathbb{R}^m): \quad A(\alpha x + \beta y) = \alpha(Ax) + \beta(Ay) \tag{5.21}$$

gilt. Wir rechnen dies komponentenweise nach. Es gilt

$$\begin{aligned}\sum_{\mu=1}^{m} a_{\nu\mu}(\alpha x_\mu + \beta y_\mu) &= \sum_{\mu=1}^{m} \alpha(a_{\nu\mu}x_\mu) + \sum_{\mu=1}^{m} \beta(a_{\nu\mu}y_\mu) \\ &= \alpha\left(\sum_{\mu=1}^{m} a_{\nu\mu}x_\mu\right) + \beta\left(\sum_{\mu=1}^{m} a_{\nu\mu}y_\mu\right).\end{aligned}$$

Nach den Definitionen (5.18) und (5.19) ist die Linearitätsbedingung (5.21) damit komponentenweise nachgerechnet. Die Aussagen (1) und (2) sind Umformulierungen der entsprechenden Aussagen des Satzes 5.8. □

Aus den Sätzen 5.8 und 5.9 folgt eine koordinatenfreie Charakterisierung der transponierten Matrix einer $(n \times m)$-Matrix mit Hilfe des euklidischen inneren Produktes.

Satz 5.10. *Sei $A \in \mathbb{R}^{n\times m}$ beliebig gegeben. Dann ist A^t die einzige Matrix $B \in \mathbb{R}^{m\times n}$, die die Bedingung*

$$(\forall x \in \mathbb{R}^m)(\forall y \in \mathbb{R}^n): \quad \langle Ax, y\rangle = \langle x, By\rangle \tag{5.22}$$

erfüllt. Auf der linken Seite der Gleichung (5.22) *ist $\langle\, .\, ,\, .\, \rangle$ das euklidische innere Produkt des $\mathbb{R}^n$. Auf der rechten Seite der Gleichung* (5.22) *ist $\langle\, .\, ,\, .\, \rangle$ das euklidische innere Produkt des $\mathbb{R}^m$.*

Beweis. Wir nehmen an, dass die Bedingung (5.22) erfüllt ist. Einsetzen der kanonischen Basisvektoren $e_\nu^{(n)}$ für $y \in \mathbb{R}^n$ respektive $e_\mu^{(m)}$ für $x \in \mathbb{R}^m$ in Formel (2) des Satzes 5.8 liefert

$$b_{\mu\nu} = \langle e_\mu^{(m)}, Be_\nu^{(n)} \rangle = \langle Ae_\mu^{(m)}, e_\nu^{(n)} \rangle = \langle e_\nu^{(n)}, Ae_\mu^{(m)} \rangle = a_{\nu\mu}$$

für alle $\nu = 1, \ldots, n$ und alle $\mu = 1, \ldots, m$. Folglich ist $B = A^t$.

Wir nehmen umgekehrt an, dass $B = A^t$ gilt. Nach Definition gilt dann $b_{\mu\nu} = a_{\nu\mu}$ für alle $\nu = 1, \ldots, n$ und alle $\mu = 1, \ldots, m$. Mit Aussage (2) des Satzes 5.8 folgt

$$\langle e_\mu^{(m)}, Be_\nu^{(n)} \rangle = b_{\mu\nu} = a_{\nu\mu} = \langle e_\nu^{(n)}, Ae_\mu^{(m)} \rangle$$

für alle $\nu = 1, \ldots, n$ und alle $\mu = 1, \ldots, m$. Weil das euklidische innere Produkt nach Satz 3.4 bilinear ist, folgt nun mit Satz 5.9 und Aussage (5) aus Satz 5.1 die Behauptung (5.22). □

Aus dem Beweis des Satzes 5.10 folgt, dass (5.22) und

$$(\forall \mu = 1, \ldots, m)(\forall \nu = 1, \ldots, n): \quad \langle Ae_\mu^{(m)}, e_\nu^{(n)} \rangle = \langle e_\mu^{(m)}, Be_\nu^{(n)} \rangle \tag{5.23}$$

äquivalente Aussagen sind.

Definition 5.11. *Eine quadratische Matrix $A = (a_{\nu\mu}) \in \mathbb{R}^{n \times n}$ heißt* symmetrisch*, wenn*

$$A = A^t \tag{5.24}$$

gilt. Die Bedingung (5.24) *ist genau dann erfüllt, wenn*

$$(\forall \mu = 1, \ldots, m)(\forall \nu = 1, \ldots, n): \quad a_{\nu\mu} = a_{\mu\nu} \tag{5.25}$$

erfüllt ist.

Aus Satz 5.10 ergibt sich eine koordinatenfreie Charakterisierung symmetrischer Matrizen.

Satz 5.12. *Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ ist genau dann eine symmetrische Matrix, wenn sie die Bedingung*

$$(\forall x, y \in \mathbb{R}^n): \quad \langle Ax, y \rangle = \langle x, Ay \rangle \tag{5.26}$$

erfüllt, wobei $\langle\, .\, ,\, .\, \rangle$ das euklidische innere Produkt des $\mathbb{R}^n$ ist.

Beispiel 5.13. Wir rechnen (5.26) für die symmetrische Matrix $A \in \mathbb{R}^{4\times 4}$ und die Spaltenvektoren x, $y \in \mathbb{R}^4$ mit

$$A = \begin{pmatrix} -7 & 0 & 2 & -2 \\ 0 & -2 & -5 & -1 \\ 2 & -5 & 1 & 5 \\ -2 & -1 & 5 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} 3 \\ -4 \\ 0 \\ 2 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 5 \end{pmatrix}$$

explizit nach.

$$Ax = \begin{pmatrix} -7 & 0 & 2 & -2 \\ 0 & -2 & -5 & -1 \\ 2 & -5 & 1 & 5 \\ -2 & -1 & 5 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -4 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -25 \\ 6 \\ 36 \\ 0 \end{pmatrix},$$

$$\langle Ax, y \rangle = \begin{pmatrix} -25 \\ 6 \\ 36 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -2 \\ 5 \end{pmatrix} = -97\,,$$

$$Ay = \begin{pmatrix} -7 & 0 & 2 & -2 \\ 0 & -2 & -5 & -1 \\ 2 & -5 & 1 & 5 \\ -2 & -1 & 5 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \\ 5 \end{pmatrix} = \begin{pmatrix} -21 \\ 5 \\ 25 \\ -7 \end{pmatrix},$$

$$\langle x, Ay \rangle = \begin{pmatrix} 3 \\ -4 \\ 0 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -21 \\ 5 \\ 25 \\ -7 \end{pmatrix} = -97\,.$$

□

Wir kommen nun zur Definition der Matrixmultiplikation. Die Definition wird so getroffen, dass die früheren Definitionen (5.18) und (5.19) im Falle des Produktes einer $(n \times m)$-Matrix mit einer $(m \times 1)$-Matrix reproduziert werden. Wieder lassen wir uns von der Auffassung leiten, dass Matrizen Zusammenstellungen ihrer Spalten sind.

Seien $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ und $B = (b_{\mu\lambda}) \in \mathbb{R}^{m\times l}$ beliebig gegeben. Sei dann $C = (c_{\nu\lambda}) \in \mathbb{R}^{n\times l}$ die Matrix mit den Einträgen

$$c_{\nu\lambda} = \sum_{\mu=1}^{m} a_{\nu\mu} b_{\mu\lambda}\,, \quad \nu = 1, \ldots, n\,, \quad \lambda = 1, \ldots, l\,. \tag{5.27}$$

Wir setzen

$$A \circ B \equiv AB \equiv C \in \mathbb{R}^{n\times l}\,. \tag{5.28}$$

Die *Matrixmultiplikation* $\circ$ wird durch

$$\circ : \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times l} \to \mathbb{R}^{n\times l}\,, \quad (A,B) \mapsto A \circ B = AB \tag{5.29}$$

definiert. Wir heben ausdrücklich hervor, dass die *Produktmatrix* AB zweier Matrizen A und B nur dann definiert ist, wenn die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt.

Satz 5.14. *Seien $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ und $B = (b_{\mu\lambda}) \in \mathbb{R}^{m\times l}$ beliebig gegeben. Dann gelten die folgenden Aussagen.*

(1) $0_{n\times m} \circ B = 0_{n\times l}$.

(2) $A \circ 0_{m\times l} = 0_{n\times l}$.

(3) $E_n \circ A = A \circ E_m = A$.

Beweis. Nach (5.27) verschwinden alle Einträge der Produktmatrix AB, wenn eine der beiden Matrizen A oder B eine Nullmatrix ist. Damit sind die Aussagen (1) und (2) bewiesen. Die Gültigkeit der Aussage (3) folgt aus

$$\sum_{\tilde{\nu}=1}^{n} \delta^{(n)}_{\nu\tilde{\nu}} a_{\tilde{\nu}\mu} = \delta^{(n)}_{\nu\nu} a_{\nu\mu} = a_{\nu\mu}\,, \quad \nu = 1,\ldots,n\,, \quad \mu = 1,\ldots,m$$

respektive aus

$$\sum_{\tilde{\mu}=1}^{m} a_{\nu\tilde{\mu}} \delta^{(m)}_{\tilde{\mu}\mu} = a_{\nu\mu}\delta^{(m)}_{\mu\mu} = a_{\nu\mu}\,, \quad \nu = 1,\ldots,n\,, \quad \mu = 1,\ldots,m\,.$$

Damit ist der Beweis beendet. □

Wie allgemein üblich schreiben wir die Formeln der Aussagen des Satzes 5.14 in der vereinfachten Form

$$0B = 0\,, \quad A0 = 0\,, \quad EA = AE = A\,.$$

Zu Beginn des vorliegenden Abschnittes haben wir eigens darauf hingewiesen, dass beim Rechnen mit Matrizen neue Phänomene auftreten: Die Matrixmultiplikation verletzt das Kommutativgesetz und lässt Nullteiler zu. Beide Phänomene treten bereits beim Rechnen mit (2×2)-Matrizen auf, wie das folgende Beispiel zeigt.

Beispiel 5.15. Die (2×2)-Matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

ist ein Beispiel für eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ mit den Eigenschaften

$$A \neq 0\,, \quad A^t A \neq A A^t\,, \quad AA = 0\,.$$

Eine einfache Rechnung ergibt

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\,,$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\,,$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\,.$$

Die Matrixmultiplikation ist in $\mathbb{R}^{2 \times 2}$ weder *kommutativ* noch *nullteilerfrei.* □

Bevor wir die Assoziativgesetze formulieren und beweisen, notieren wir die Regeln über das Herausziehen von Skalaren.

Satz 5.16. *Für $A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{m \times l}$, $x \in \mathbb{R}^m$ und $\alpha \in \mathbb{R}$ gelten:*

(1) $A(\alpha x) = (\alpha A)x = \alpha(Ax)$.

(2) $A(\alpha B) = (\alpha A)B = \alpha(AB)$.

Nun wenden wir uns den Assoziativgesetzen zu. Eine wichtige Anwendung des Assoziativgesetzes (5) des folgenden Satzes 5.17 ist der Nachweis der Einzigkeit der inversen Matrix in Satz 5.21.

Satz 5.17. *Es seien $A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{m \times l}$, $C \in \mathbb{R}^{l \times k}$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^l$ sowie α, $\beta \in \mathbb{R}$ beliebig gegeben. Dann gelten die folgenden Assoziativgesetze.*

(1) $(\alpha\beta)x = \alpha(\beta x)$.

(2) $(\alpha\beta)A = \alpha(\beta A)$.

(3) $\alpha(Ax) = (\alpha A)x$.

(4) $A(By) = (AB)y$.

(5) $A(BC) = (AB)C$.

Beweis. Nachweis von (3). Für beliebige Skalare $\alpha \in \mathbb{R}$ und Matrizen $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ sowie Spaltenvektoren $x = (x_\mu) \in \mathbb{R}^m$ gilt

$$\alpha \left(\sum_{\mu=1}^{m} a_{\nu\mu} x_\mu \right) = \sum_{\mu=1}^{m} (\alpha a_{\nu\mu}) x_\mu$$

für alle Indizes $\nu = 1, \dots, n$.

Nachweis von (4). Für beliebige Matrizen $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ und $B = (b_{\mu\lambda}) \in \mathbb{R}^{m\times l}$ sowie Spaltenvektoren $x = (x_\lambda) \in \mathbb{R}^l$) gilt

$$\begin{aligned}\sum_{\mu=1}^{m} a_{\nu\mu} \left(\sum_{\lambda=1}^{l} b_{\mu\lambda} x_\lambda \right) &= \sum_{\mu=1}^{m} \left(\sum_{\lambda=1}^{l} a_{\nu\mu} (b_{\mu\lambda} x_\lambda) \right) \\ &= \sum_{\lambda=1}^{l} \left(\left(\sum_{\mu=1}^{m} a_{\nu\mu} b_{\mu\lambda} \right) x_\lambda \right)\end{aligned}$$

für alle Indizes $\nu = 1, \dots, n$.

Nachweis von (5). Für beliebige Matrizen $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$ und $B = (b_{\mu\lambda}) \in \mathbb{R}^{m\times l}$ sowie $C = (c_{\lambda\kappa}) \in \mathbb{R}^{l\times k}$) gilt

$$\begin{aligned}\sum_{\mu=1}^{m} a_{\nu\mu} \left(\sum_{\lambda=1}^{l} b_{\mu\lambda} c_{\lambda\kappa} \right) &= \sum_{\mu=1}^{m} \left(\sum_{\lambda=1}^{l} a_{\nu\mu} (b_{\mu\lambda} c_{\lambda\kappa}) \right) \\ &= \sum_{\lambda=1}^{l} \left(\left(\sum_{\mu=1}^{m} a_{\nu\mu} b_{\mu\lambda} \right) c_{\lambda\kappa} \right)\end{aligned}$$

für alle Indizes $\nu = 1, \dots, n$ und $\kappa = 1, \dots, k$. □

Wir studieren das Verhalten von Matrizenprodukten bei Transposition. Wie bei der Inversion wird die Reihenfolge der Matrixmultiplikationen umgekehrt. Vergleiche die Sätze 5.18 und 5.23.

Satz 5.18. *Es gelten die folgenden Aussagen.*

(1) $(\forall A \in \mathbb{R}^{n\times m})(\forall B \in \mathbb{R}^{m\times l}): \quad (AB)^t = B^t A^t.$

(2) *Sei $A \in \mathbb{R}^{n\times m}$ beliebig gegeben. Dann gilt $A^t \in \mathbb{R}^{m\times n}$. Außerdem sind die beiden Produktmatrizen*

$$A^t A \in \mathbb{R}^{m\times m}, \quad AA^t \in \mathbb{R}^{n\times n}$$

symmetrische Matrizen.

Beweis. Nachweis von (1). Ein direkter Beweis kann durch explizites Nachrechnen der Gleichheit der Matrixeinträge erbracht werden. Wir geben stattdessen einen koordinatenfreien Beweis mit Hilfe des Satzes 5.10 und des Assoziativgesetzes (4) aus Satz 5.17. Für alle $x \in \mathbb{R}^l$ und alle $y \in \mathbb{R}^n$ gilt

$$\begin{aligned}\langle x, (AB)^t y\rangle &= \langle (AB)x, y\rangle = \langle A(Bx), y\rangle \\ &= \langle Bx, A^t y\rangle = \langle x, B^t(A^t y)\rangle = \langle x, (B^t A^t)y\rangle\,.\end{aligned}$$

Aus Satz 5.10 folgt $(AB)^t = B^t A^t$.

Nachweis von (2). Nach Definition der Matrixmultiplikation gilt $A^t A \in \mathbb{R}^{m\times m}$ und $AA^t \in \mathbb{R}^{n\times n}$. Aussage (1) liefert die Symmetrie der beiden Produktmatrizen:

$$(A^t A)^t = A^t A^{tt} = A^t A\,, \quad (AA^t)^t = A^{tt} A^t = AA^t\,.$$

Damit ist der Beweis beendet. □

Beispiel 5.19.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{pmatrix}\,, \quad A^t = \begin{pmatrix} 1 & 6 \\ 2 & 5 \\ 3 & 4 \end{pmatrix}\,.$$

$$B = \begin{pmatrix} 1 & 0 & -2 & -3 \\ 4 & 0 & 5 & -6 \\ 1 & -1 & 3 & -2 \end{pmatrix}\,, \quad B^t = \begin{pmatrix} 1 & 4 & 1 \\ 0 & 0 & -1 \\ -2 & 5 & 3 \\ -3 & -6 & -2 \end{pmatrix}\,.$$

$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 & -3 \\ 4 & 0 & 5 & -6 \\ 1 & -1 & 3 & -2 \end{pmatrix} = \begin{pmatrix} 12 & -3 & 17 & -21 \\ 30 & -4 & 25 & -56 \end{pmatrix}\,.$$

$$B^t A^t = \begin{pmatrix} 1 & 4 & 1 \\ 0 & 0 & -1 \\ -2 & 5 & 3 \\ -3 & -6 & -2 \end{pmatrix} \begin{pmatrix} 1 & 6 \\ 2 & 5 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 12 & 30 \\ -3 & -4 \\ 17 & 25 \\ -21 & -56 \end{pmatrix}\,.$$

□

Beispiel 5.20.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{pmatrix}\,, \quad A^t = \begin{pmatrix} 1 & 6 \\ 2 & 5 \\ 3 & 4 \end{pmatrix}\,,$$

$$A^tA = \begin{pmatrix} 37 & 32 & 27 \\ 32 & 29 & 26 \\ 27 & 26 & 25 \end{pmatrix}, \quad AA^t = \begin{pmatrix} 14 & 28 \\ 28 & 77 \end{pmatrix}.$$

□

Wir betrachten nun eine besondere Klasse von quadratischen Matrizen.

Satz und Definition 5.21. *Sei $A \in \mathbb{R}^{n\times n}$ beliebig gegeben. Dann sind die beiden folgenden Aussagen* (1) *und* (2) *äquivalent.*

(1) $(\exists B \in \mathbb{R}^{n\times n}): \quad AB = BA = E.$

(2) $(\exists!\, B \in \mathbb{R}^{n\times n}): \quad AB = BA = E.$

Wenn eine der äquivalenten Bedingungen (1) *oder* (2) *erfüllt ist, dann heißt die quadratische Matrix A* invertierbar*. Andernfalls heißt A* singulär*. Beide Begriffe werden nur für quadratische Matrizen definiert.*

Beweis. Es ist klar, dass (1) aus (2) folgt. Wir zeigen umgekehrt, dass (2) aus (1) folgt. Es ist lediglich die Einzigkeit zu beweisen. Die Existenz wird in (1) bereits vorausgesetzt. Seien Matrizen $B, C \in \mathbb{R}^{n\times n}$ mit

$$AB = BA = E\,, \quad AC = CA = E$$

gegeben. Dann folgt mit dem Assoziativgesetz (5) aus Satz 5.17 und der Eigenschaft (3) der Einheitsmatrix aus Satz 5.14, dass

$$B = BE = B(AC) = (BA)C = EC = C$$

gilt. Dies liefert die Einzigkeitsaussage. Damit ist der Beweis beendet. □

Definition 5.22. *Wenn die Matrix $A \in \mathbb{R}^{n\times n}$ invertierbar ist, dann heißt die eindeutig bestimmte Matrix $B \in \mathbb{R}^{n\times n}$ mit $AB = BA = E$ die* inverse Matrix *oder die* Inverse *zu A. Die inverse Matrix zu A wird mit A^{-1} bezeichnet. Sei*

$$\mathrm{GL}(n,\mathbb{R}) = \{\, A \in \mathbb{R}^{n\times n} \mid (\exists B \in \mathbb{R}^{n\times n}): \quad AB = BA = E \,\}$$

die Menge der invertierbaren reellen $(n \times n)$-Matrizen. Siehe Satz 5.24*. Wie üblich identifizieren wir $\mathbb{R}^{1\times 1}$ und $\mathbb{R}$ vermittels der Abbildung $(\alpha) \mapsto \alpha$. Wir setzen*

$$\mathbb{R}^\times = \mathbb{R} \setminus \{0\} = \{\, \alpha \in \mathbb{R} \mid \alpha \neq 0 \,\}.$$

Es gilt $\mathbb{R}^\times = \mathrm{GL}(1,\mathbb{R})$. Eine reelle Zahl $\alpha \in \mathbb{R}^\times$ heißt invertierbar*. Die Invertierbarkeit einer Matrix und die Inverse einer Matrix können mit dem Verfahren von Gauß-Jordan bestimmt werden. Siehe Satz* 6.13*.*

Wir untersuchen das Verhalten von Matrizenprodukten bei Inversion. Wie bei der Transposition wird die Reihenfolge der Matrixmultiplikationen umgekehrt. Inversion und Transposition sind auf $\mathrm{GL}(n,\mathbb{R})$ vertauschbare Operationen.

Satz 5.23. *Es gelten die folgenden Aussagen.*

(1) *Es gilt* $E \in \mathrm{GL}(n,\mathbb{R})$ *mit* $E^{-1} = E$.

(2) $(\forall \alpha \in \mathbb{R}^\times)(\forall A \in \mathrm{GL}(n,\mathbb{R})): \quad \alpha A \in \mathrm{GL}(n,\mathbb{R})\,, \quad (\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$.

(3) $(\forall A, B \in \mathrm{GL}(n,\mathbb{R})): \quad AB \in \mathrm{GL}(n,\mathbb{R})\,, \quad (AB)^{-1} = B^{-1}A^{-1}$.

(4) $(\forall A \in \mathrm{GL}(n,\mathbb{R})): \quad A^t \in \mathrm{GL}(n,\mathbb{R})\,, \quad (A^t)^{-1} = (A^{-1})^t$.

Beweis. Nachweis von (1). Es gilt

$$EE = E\,.$$

Damit ist (1) bewiesen. Aus den Regeln über das Herausziehen von Skalaren und dem Assoziativgesetz für das Rechen mit Skalaren und Matrizen folgen die Beziehungen

$$\begin{aligned}(\alpha A)(\tfrac{1}{\alpha}A^{-1}) &= \alpha(A(\tfrac{1}{\alpha}A^{-1})) = \alpha(\tfrac{1}{\alpha}(AA^{-1})) \\ &= (\alpha\tfrac{1}{\alpha})(AA^{-1}) = 1 \cdot E = E\,, \\[1ex] (\tfrac{1}{\alpha}A^{-1})(\alpha A) &= \tfrac{1}{\alpha}(A^{-1}(\alpha A)) = \tfrac{1}{\alpha}(\alpha(A^{-1}A)) \\ &= (\tfrac{1}{\alpha}\alpha)(A^{-1}A) = 1 \cdot E = E\,.\end{aligned}$$

Damit ist (2) bewiesen. Aus dem Assoziativgesetz für die Matrixmultiplikation folgen die Beziehungen

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(B(B^{-1}A^{-1})) = A((BB^{-1})A^{-1}) \\ &= A(EA^{-1}) = AA^{-1} = E\,, \\[1ex] (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}(AB)) = B^{-1}((A^{-1}A)B) \\ &= B^{-1}(EB) = B^{-1}B = E\,.\end{aligned}$$

Damit ist (3) bewiesen. Aus der Rechenregel über die Transposition eines Matrixproduktes folgen die Beziehungen

$$\begin{aligned}(A^{-1})^t A^t &= (AA^{-1})^t = E^t = E\,, \\ A^t(A^{-1})^t &= (A^{-1}A)^t = E^t = E\,.\end{aligned}$$

Damit ist auch (4) bewiesen. □

Die invertierbaren $(n \times n)$-Matrizen bilden mit der Komposition von Matrizen eine Gruppe im Sinne von Definition 21.1.

Satz 5.24. *Sei $n \in \mathbb{N}$. Das Paar $(\mathrm{GL}(n, \mathbb{R}), \circ)$ ist eine Gruppe. Dabei ist die Einheitsmatrix E_n das neutrale Element dieser Gruppe.*

Nun definieren wir die Summe von Matrizen. Seien $A = (a_{\nu\mu}) \in \mathbb{R}^{n \times m}$ und $B = (b_{\nu\mu}) \in \mathbb{R}^{n \times m}$ beliebig gegeben. Sei dann $C = (c_{\nu\mu}) \in \mathbb{R}^{n \times m}$ die Matrix mit den Einträgen

$$c_{\nu\mu} = a_{\nu\mu} + b_{\nu\mu}\,, \quad \nu = 1, \ldots, n\,, \quad \mu = 1, \ldots, m\,. \tag{5.30}$$

Wir setzen

$$A + B = C\,. \tag{5.31}$$

Die Matrix C heißt *Summe der Matrizen A und B*. Zusätzlich wird

$$-B = (-1)B\,, \quad A - B = A + (-B) \tag{5.32}$$

gesetzt. Wir heben ausdrücklich hervor, dass die Summe zweier Matrizen A und B nur dann erklärt ist, wenn erstens die Zeilenzahl von A mit Zeilenzahl von B und zweitens die Spaltenzahl von A mit der Spaltenzahl von B übereinstimmt.

Satz 5.25. *Seien n, m, $k \in \mathbb{N}$. Dann gelten:*

(1) *Das Tripel $(\mathbb{R}^{n \times m}, \cdot, +)$ mit den durch* (5.15), (5.30), (5.31) *definierten Operationen ist ein reeller Vektorraum.*

(2) *Für $i = 1, \ldots, n$ und $j = 1, \ldots m$ sei $E_{ij} = E_{ij}^{(nm)} \in \mathbb{R}^{n \times m}$ die Matrix, die an der Stelle (i, j) den Eintrag Eins und sonst nur verschwindende Einträge besitzt. Es gilt*

$$A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} E_{ij}^{(nm)} \right)$$

für alle $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. Die Menge

$$\mathcal{E}_{nm} = \{E_{ij} \in \mathbb{R}^{n \times m} \mid i = 1, \ldots, n\,,\ j = 1, \ldots m\}$$

ist eine Basis von $\mathbb{R}^{n \times m}$.

(3) $\dim(\mathbb{R}^{n \times m}) = nm\,.$

(4) *Sei $C \in \mathbb{R}^{m \times k}$ gegeben. Dann ist $\rho_C \colon \mathbb{R}^{n \times m} \to \mathbb{R}^{n \times k}$ mit $A \mapsto AC$ linear.*

(5) *Sei $C \in \mathbb{R}^{k \times n}$ gegeben. Dann ist $\lambda_C \colon \mathbb{R}^{n \times m} \to \mathbb{R}^{k \times m}$ mit $A \mapsto CA$ linear.*

(6) *Die Transposition ${}^t \colon \mathbb{R}^{n \times n} \to \mathbb{R}^{m \times n}$ mit $A \mapsto A^t$ ist ein Isomorphismus.*

6 Kern und Bild einer Matrix. Gauß-Jordan

Eine Matrix $A \in \mathbb{R}^{n\times m}$ fassen wir als lineare Abbbildung $\varphi_A\colon \mathbb{R}^m \to \mathbb{R}^n$ mit $x \mapsto Ax$ auf. Den Kern und das Bild von φ_A bezeichnen wir als Kern respektive Bild der Matrix A. Wir setzen also

$$\ker(A) = \ker(\varphi_A)\,, \quad \mathrm{im}(A) = \mathrm{im}(\varphi_A)\,.$$

Wir schreiben die Definitionen von $\ker(A)$ und $\mathrm{im}(A)$ im Hinblick auf die Gleichungstheorie aus und formulieren erste Eigenschaften. Siehe 2.8 und 23.3.

Satz und Definition 6.1. *Seien $n, m \in \mathbb{N}$. Sei $A = (a_1, \dots, a_m) \in \mathbb{R}^{n\times m}$ eine $(n \times m)$-Matrix mit den Spalten $a_1, \dots, a_m \in \mathbb{R}^n$.*

(1) *Die Teilmenge*

$$\ker(A) = \{\, x \in \mathbb{R}^m \mid Ax = 0 \,\}$$

heißt Kern *oder* Nullraum *von A. Weil $x \mapsto Ax$ nach Satz 5.9 eine lineare Abbildung φ_A von $\mathbb{R}^m$ nach $\mathbb{R}^n$ ist, ist $\ker(A) \subseteq \mathbb{R}^m$ ein Teilvektorraum des $\mathbb{R}^m$. Der Teilvektorraum $\ker(A)$ besteht aus allen Lösungen $x \in \mathbb{R}^m$ des homogenen Gleichungssystems $Ax = 0$.*

(2) *Die Teilmenge*

$$\mathrm{im}(A) = \{\, y \in \mathbb{R}^n \mid (\exists x \in \mathbb{R}^m):\; Ax = y \,\}$$

heißt Bild *oder* Spaltenraum *von A. Weil $\varphi_A\colon x \mapsto Ax$ nach Satz 5.9 eine lineare Abbildung von $\mathbb{R}^m$ nach $\mathbb{R}^n$ ist, ist $\mathrm{im}(A) \subseteq \mathbb{R}^n$ ein Teilvektorraum des $\mathbb{R}^n$. Die Menge $\mathrm{im}(A)$ besteht aus allen Bildern Ax mit $x \in \mathbb{R}^m$. Aus Satz 5.8, Aussage (1) folgt, dass*

$$(\forall \mu = 1, \dots, m): \quad a_\mu \in \mathrm{im}(A)$$

gilt. Dies rechtfertigt die Bezeichnung Spaltenraum für den Teilvektorraum $\mathrm{im}(A)$. Ein inhomogenes Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $b \in \mathrm{im}(A)$ gilt.

(3) *φ_A ist genau dann injektiv, wenn $\ker(A) = \{0\}$ gilt.*

(4) *φ_A ist genau dann surjektiv, wenn $\mathrm{im}(A) = \mathbb{R}^n$ gilt.*

Siehe die Sätze 7.6 und 13.1.

Der Kern und das Bild einer Matrix sind endlich-dimensionale Teilvektorräume. Diese lassen sich durch Basen bequem beschreiben. In den Sätzen 6.17 und 6.18 geben wir Verfahren zur Bestimmung von entsprechenden Basen an. Daraus ergeben sich ein Lösbarkeitskriterium sowie ein Lösungsverfahren für lineare Gleichungssysteme in den Sätzen 6.19 und 6.20. Ein zentrales Hilfsmittel ist das Gauß-Jordan-Verfahren.

Wir werden in Satz 6.8 sehen, dass der Kern einer Matrix unter elementaren Zeilenumformungen invariant ist. Zum Nachweis überlegen wir uns, dass die Anwendung einer elementaren Zeilenumformung zu der linksseitigen Multiplikation mit einer sogenannten *Elementarmatrix* äquivalent ist.

Zuerst einmal übersetzen wir die elementaren Zeilenumformungen in die Sprache der Matrixmultiplikationen. Sei

$$A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{R}^{n \times m}$$

eine $(n \times m)$-Matrix mit den Zeilen $z_1^t, \ldots, z_n^t \in \mathbb{R}^m$. Im Folgenden bezeichnen wir mit $i,\ j \in \{1, \ldots, n\}$ die Zeilenindizes. Wir listen die *elementaren Zeilenumformungen* auf. Es gibt drei Typen. Elementare Zeilenumformungen werden eingeführt, um äquivalente Umformungen von linearen Gleichungssystemen vornehmen zu können. Im Matrixkalkül bedeutet dies, dass sie durch invertierbare $(n \times n)$-Matrizen repräsentiert werden.

(1) Die Vertauschung $Z_{i \leftrightarrow j}$ der i-ten mit der j-ten Zeile wird durch die linksseitige Multiplikation von A mit der Matrix

$$E_{i \leftrightarrow j} \in \mathrm{GL}(n, \mathbb{R}) \tag{6.1}$$

erreicht, wobei $E_{i \leftrightarrow j}$ aus der Einheitsmatrix $E \in \mathbb{R}^{n \times n}$ durch Vertauschen der i-ten und der j-ten Zeile entsteht. Die Matrix $E_{i \leftrightarrow j}$ ist involutorisch, das heißt, es gilt

$$E_{i \leftrightarrow j} E_{i \leftrightarrow j} = E\,. \tag{6.2}$$

Die Matrix $E_{i \leftrightarrow j}$ ist invertierbar mit

$$E_{i \leftrightarrow j}^{-1} \equiv (E_{i \leftrightarrow j})^{-1} = E_{i \leftrightarrow j}\,. \tag{6.3}$$

Im Fall $i = j$ ist $E_{i \leftrightarrow j}$ die Einheitsmatrix E.

(2) Die Multiplikation αZ_i der i-ten Zeile mit $\alpha \in \mathbb{R}^\times$ wird durch die linksseitige Multiplikation von A mit der Matrix

$$E_{\alpha;i} \in \mathrm{GL}(n, \mathbb{R}) \tag{6.4}$$

erreicht. Die Matrix $E_{\alpha;i}$ entsteht aus der Einheitsmatrix $E \in \mathbb{R}^{n \times n}$, indem die i-te Zeile e_i^t durch die Zeile αe_i^t ersetzt wird. Es gilt

$$E_{\alpha;i}^{-1} \equiv (E_{\alpha;i})^{-1} = E_{\frac{1}{\alpha};i}\,. \tag{6.5}$$

Im Fall $\alpha = 1$ ist $E_{\alpha;i}$ die Einheitsmatrix.

(3) Die Addition $Z_j + \alpha Z_i$ der α-fachen i-ten Zeile zur j-Zeile mit $\alpha \in \mathbb{R}$ und $i \neq j$ wird durch die linksseitige Multiplikation von A mit der Matrix

$$E_{j;\alpha;i} \in \mathrm{GL}(n, \mathbb{R}) \tag{6.6}$$

erreicht. Die Matrix $E_{j;\alpha;i}$ entsteht aus der Einheitsmatrix $E \in \mathbb{R}^{n\times n}$, indem die j-te Zeile e_j^t durch die Zeile $e_j^t + \alpha e_i^t$ ersetzt wird. Es gilt

$$E_{j;\alpha;i}^{-1} \equiv (E_{j;\alpha;i})^{-1} = E_{j;-\alpha;i}\,. \tag{6.7}$$

Im Fall $\alpha = 0$ ist $E_{j;\alpha;i}$ die Einheitsmatrix.

Um hervorzuheben, dass die betrachteten Elementarmatrizen $(n \times n)$-Matrizen sind, schreiben wir $E_{i\leftrightarrow j}^{(n)}$, $E_{\alpha;i}^{(n)}$ oder $E_{j;\alpha;i}^{(n)}$ anstelle von $E_{i\leftrightarrow j}$, $E_{\alpha;i}$ oder $E_{j;\alpha;i}$.

Satz und Definition 6.2. *Seien $n, m \in \mathbb{N}$. Die Anwendung elementarer Zeilenumformungen einer Matrix $A \in \mathbb{R}^{n\times m}$ kann durch die linksseitige Multiplikation mit invertierbaren Matrizen vom Typ*

(1) $E_{i\leftrightarrow j}$,

(2) $E_{\alpha;i}$ *mit* $\alpha \in \mathbb{R}\setminus\{0\}$ *oder*

(3) $E_{j;\alpha;i}$ *mit* $\alpha \in \mathbb{R}$ *und* $i \neq j$

beschrieben werden. Wir nennen diese Matrizen Elementarmatrizen *und bezeichnen mit* $\mathrm{Z}(n,\mathbb{R})$ *die Menge aller Elementarmatrizen aus* $\mathrm{GL}(n,\mathbb{R})$. *Weiter bestehe die Teilmenge $\tilde{Z}(n,\mathbb{R}) \subseteq \mathrm{GL}(n,\mathbb{R})$ aus allen endlichen Produkten von Matrizen aus $Z(n,\mathbb{R})$.*

Sei $A \in \mathbb{R}^{n\times m}$. Die Wirkung der linkseitigen Multiplikation mit einer der Elementarmatrizen $E_{i\leftrightarrow j}$, $E_{\alpha;i}$, $E_{j;\alpha;i} \in \mathrm{GL}(n,\mathbb{R})$ schreiben wir schematisch in der Form

$$\begin{array}{l|l} A & \\ \hline E_{i\leftrightarrow j}\circ A & Z_{i\leftrightarrow j} \end{array}\,, \qquad \begin{array}{l|l} A & \\ \hline E_{\alpha;i}\circ A & \alpha Z_i \end{array}\,, \qquad \begin{array}{l|l} A & \\ \hline E_{j;\alpha;i}\circ A & Z_j + \alpha Z_i \end{array}\,.$$

Dabei ist jeweils die rechte Spalte der drei Schemata die Kommentarspalte. Insbesondere für $A = E = E_n$ erhalten wir

$$\begin{array}{l|l} E & \\ \hline E_{i\leftrightarrow j} & Z_{i\leftrightarrow j} \end{array}\,, \qquad \begin{array}{l|l} E & \\ \hline E_{\alpha;i} & \alpha Z_i \end{array}\,, \qquad \begin{array}{l|l} E & \\ \hline E_{j;\alpha;i} & Z_j + \alpha Z_i \end{array}\,.$$

Wir weisen auf die entsprechende Erörterung der Wirkung von Permutationsmatrizen im Anschluss an Beispiel 9.9 hin.

Beispiel 6.3. Wir notieren einige Elementarmatrizen explizit:

$$E_{2\leftrightarrow 4}^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad E_{2\leftrightarrow 4}^{(5)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$E_{7;3}^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad E_{7;3}^{(5)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} ,$$

$$E_{3;17;1}^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 17 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad E_{3;17;1}^{(5)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 17 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} .$$

Die Wirkung von $E_{3;17;1}^{(4)}$ beschreiben wir in Beispiel 6.4. □

Beispiel 6.4. Gegeben sei $A \in \mathbb{R}^{4\times 5}$ mit

$$A = \begin{pmatrix} 1 & 2 & 2 & 4 & 3 \\ 2 & 4 & 3 & 8 & 4 \\ 3 & 6 & 4 & 12 & 5 \\ 4 & 8 & 5 & 16 & 4 \end{pmatrix}$$

Die linksseitige Multiplikation mit der Elementarmatrix $E_{3;17;1}^{(4)}$ ergibt

$$E_{3;17;1}^{(4)} \circ A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 17 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 2 & 4 & 3 \\ 2 & 4 & 3 & 8 & 4 \\ 3 & 6 & 4 & 12 & 5 \\ 4 & 8 & 5 & 16 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 2 & 4 & 3 \\ 2 & 4 & 3 & 8 & 4 \\ 20 & 40 & 38 & 80 & 56 \\ 4 & 8 & 5 & 16 & 4 \end{pmatrix} .$$

Schematisch können wir dies folgendermaßen beschreiben:

$$\begin{array}{rrrrr|l} 1 & 2 & 2 & 4 & 3 & \\ 2 & 4 & 3 & 8 & 4 & \\ 3 & 6 & 4 & 12 & 5 & \\ 4 & 8 & 5 & 16 & 4 & \\ \hline 1 & 2 & 2 & 4 & 3 & Z_1 \\ 2 & 4 & 3 & 8 & 4 & Z_2 \\ 20 & 40 & 38 & 80 & 56 & Z_3 + 17Z_1 \\ 4 & 8 & 5 & 16 & 4 & Z_4 \end{array}$$

□

Mit Hilfe von elementaren Zeilenumformungen kann eine beliebige Matrix auf *obere Stufenform* gebracht werden. Siehe Satz 6.6. Eine obere Stufenform hängt allerdings von den verwendeten Zeilenumformungen ab. Die *Treppennormalform* hängt nach Satz 6.13 dagegen *nicht* von der Wahl der elementaren Zeilenumformungen ab.

Definition 6.5. *Seien $n,\, m \in \mathbb{N}$.*

(1) *Wenn in einer Zeile einer Matrix Einträge aus $\mathbb{R}^{\times}$ vorkommen, so nennen wir den Eintrag aus $\mathbb{R}^{\times}$ mit dem kleinsten Spaltenindex das* Pivotelement *dieser Zeile. Eine Nullzeile besitzt demnach kein Pivotelement. Eine Matrix, die nicht die Nullmatrix ist, heißt* pivotnormiert*, wenn alle Pivotelemente gleich* 1 *sind. Die Einheitsmatrix E_n ist pivotnormiert.*

(2) *Sei $A \in \mathbb{R}^{n\times m}$. Mit $PV(A)$ bezeichnen wir die Menge der indizierten Pivotelemente von A. Die Pivotelemente einer Matrix werden durch ihre Positionen unterschieden. Demnach besteht $PV(E_n)$ der Einheitsmatrix E_n aus n Elementen. Es gilt $PV(0_{n\times m}) = \emptyset$.*

(3) *Eine Matrix $U \in \mathbb{R}^{n\times m}$ heißt eine* obere Stufenmatrix*, wenn das Pivotelement jeder Zeile einen kleineren Spaltenindex als das Pivotelement der nächsten Zeile besitzt.*

(4) $\mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m})$ *bezeichne die Menge der oberen Stufenmatrizen aus $\mathbb{R}^{n\times m}$. Insbesondere gilt $0_{n\times m} \in \mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m})$.*

(5) *Eine Spalte einer oberen Stufenmatrix $U \in \mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m})$ heißt eine* Pivotspalte*, wenn sie ein Pivotelement enthält. Die Nullmatrix $0_{n\times m}$ enthält keine Pivotspalte.*

(6) *Wenn $U \in \mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m})$ durch elementare Zeilenumformungen aus einer Matrix $A \in \mathbb{R}^{n\times m}$ hervorgeht, dann heißt U eine* Stufenform *von A. Wir sagen dann auch, dass A auf Stufenform gebracht worden ist.*

(7) *Eine Stufenmatrix $U \in \mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m})$ heißt eine* Treppennormalform*, wenn sie eine der folgenden Bedingungen* (7.1) *oder* (7.2) *erfüllt.*

  (7.1) $U = 0$.

  (7.2) *U ist pivotnormiert und alle Einträge oberhalb der Pivotelemente verschwinden.*

(8) $\mathrm{T}(\mathbb{R}^{n\times m})$ *bezeichne die Menge der Treppennormalformen in $\mathbb{R}^{n\times m}$.*

(9) *Wenn $T \in \mathrm{T}(\mathbb{R}^{n\times m})$ aus einer Matrix $A \in \mathbb{R}^{n\times m}$ durch elementare Zeilenumformungen hervorgeht, dann heißt T eine* Treppennormalform *von A. Wir sagen dann, dass A auf Treppennormalform gebracht worden ist.*

(10) *Die* Pivotspalten *eine beliebigen Matrix $A \in \mathbb{R}^{n\times m}$ können erst in Satz und Definition* 6.13 *definiert werden.*

Wir deuten die Verfahren von Gauß und Gauß-Jordan als Satz über eine multiplikative Zerlegung von Matrizen. Die Verfahren werden im Beweis des folgenden Satzes 6.6 beschrieben. Siehe Definition 6.7.

Satz 6.6. *Seien n, $m \in \mathbb{N}$. Dann gelten:*

(1) $(\forall A \in \mathbb{R}^{n\times m})(\exists C \in \mathrm{GL}(n,\mathbb{R})): \; CA \in \mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m})$.

(2) $(\forall A \in \mathbb{R}^{n\times m})(\exists C \in \mathrm{GL}(n,\mathbb{R})): \; CA \in \mathrm{T}(\mathbb{R}^{n\times m})$.

Beweis. Ein Beweis liefern die Verfahren von Gauß respektive Gauß-Jordan. Nach diesen Verfahren können die Matrizen C als endliche Produkte von Elementarmatrizen gewählt werden. Sei $A \in \mathbb{R}^{n\times m}$. Wir setzen $k_0 = min\{n, m\}$.

Nachweis von (1). Wir beschreiben das strikte Gauß-Verfahren.

Erster Schritt. Wir konstruieren eine Matrix $C_1 \in \mathrm{GL}(n,\mathbb{R})$ derart, dass in der ersten Spalte von

$$A_1 = C_1 A \in \mathbb{R}^{n\times m}$$

höchstens der erste Eintrag von Null verschieden ist.

- Im Fall $a_1 = 0_n \in \mathbb{R}^n$ setzen wir $C_1 = E_n$.
- Im Fall $a_1 \neq 0_n \in \mathbb{R}^n$ gibt es einen kleinsten Zeilenindex $\nu \in 1, \ldots, n$ mit $a_{\nu 1} \neq 0$. Dieser Index sei ν_1. Wir vertauschen die erste und die ν_1-te Zeile von A. Die Matrix $E^{(n)}_{1\leftrightarrow\nu_1} \circ A$ besitzt an der Position $(1,1)$ den Eintrag $a_{\nu_1 1} \neq 0$. Anschließend annullieren wir die Einträge an den Positionen $(1,2)$ bis $(1,n)$ mit Hilfe der ersten Zeile. Wir setzen

  $$C_1 = E^{(n)}_{n;-a_{n1}a^{-1}_{\nu_1 1};1} \circ \ldots \circ E^{(n)}_{2;-a_{21}a^{-1}_{\nu_1 1};1} \circ E^{(n)}_{1\leftrightarrow\nu_1} \in \mathrm{GL}(n,\mathbb{R})\,.$$

  Dabei verwenden wir die Konvention, dass ein Produkt über eine leere Indexmenge gleich der Einheitsmatrix ist. Dann ist $A_1 = C_1 A \in \mathbb{R}^{n\times m}$ eine Matrix, die in ersten Spalte nur an der Position $(1,1)$ einen von Null verschiedenen Eintrag besitzt.

Zweiter Schritt. Im Fall $n, m \geq 2$ streichen wir die erste Zeile und die Spalte von A_1 und erhalten eine Matrix

$$A'_1 \in \mathbb{R}^{(n-1)\times(m-1)}\,.$$

Dann wenden wir auf A'_1 das Verfahren aus dem ersten Schritt an. Wir erhalten eine invertierbare Matrix $C'_2 \in \mathrm{GL}(n-1,\mathbb{R})$ derart, dass die erste Spalte von

$$A'_2 = C'_2 A'_1 \in \mathbb{R}^{(n-1)\times(m-1)}$$

höchstens an der Position $(1,1)$ einen von Null verschiedenen Eintrag besitzt. Wir setzen

$$C_2 = \begin{pmatrix} 1 & 0 \\ 0 & C'_2 \end{pmatrix} \in \mathrm{GL}(n,\mathbb{R})\,, \quad A_2 = C_2 A_1 = C_2 C_1 A\,.$$

Dritter Schritt. Im Fall $n, m \geq 3$ streichen wir die ersten beiden Zeilen und Spalten von A_2 und erhalten eine Matrix

$$A_2' \in \mathbb{R}^{(n-2)\times(m-2)} .$$

Wir wenden auf A_2' das Verfahren aus dem ersten Schritt an und erhalten $C_3' \in \mathrm{GL}(n-2, \mathbb{R})$ derart, dass die erste Spalte von

$$C_3' A_2' \in \mathbb{R}^{(n-2)\times(m-2)}$$

höchstens an der Position $(1, 1)$ einen von Null verschiedenen Eintrag besitzt. Wir setzen

$$C_3 = \begin{pmatrix} E_2 & 0 \\ 0 & C_3' \end{pmatrix} \in \mathrm{GL}(n, \mathbb{R}) , \quad A_3 = C_3 A_2 = C_3 C_2 C_1 A .$$

Der k-te Schritt. Im Fall

$$2 \leq k \leq k_0 = \min\{n, m\}$$

streichen wir die ersten $k-1$ Spalten und Zeilen der bereits konstruierten Matrix $A_{k-1} \in \mathbb{R}^{n\times m}$ und erhalten

$$A_k' \in \mathbb{R}^{(n-(k-1))\times(m-(k-1))} .$$

Nach dem Verfahren aus dem ersten Schritt konstruieren wir $C_k' \in \mathrm{GL}(n-2, \mathbb{R})$ derart, dass

$$C_k' A_k' \in \mathbb{R}^{(n-(k-1))\times(m-(k-1))}$$

höchstens an der Position $(1, 1)$ einen von Null verschiedenen Eintrag besitzt. Wir setzen

$$C_k = \begin{pmatrix} E_{k-1} & 0 \\ 0 & C_{k-1}' \end{pmatrix} \in \mathrm{GL}(n, \mathbb{R}) , \quad A_k = C_k A_{k-1} .$$

Der k_0-te Schritt. Wir erhalten $C_1, \ldots, C_{k_0} \in \mathrm{GL}(n, \mathbb{R})$ mit

$$A_{k_0} = C_{k_0} \circ \ldots \circ C_1 \circ A \in \mathrm{U}^{\urcorner}(\mathbb{R}^{n\times m}) .$$

Damit ist der Beweis von Aussage (1) beendet.

Nachweis von (2). Wir beschreiben das strikte Verfahren von Gauß-Jordan.

Erster Schritt. Wir konstruieren eine Matrix $C_1 \in \mathrm{GL}(n, \mathbb{R})$ derart, dass in der ersten Spalte von

$$A_1 = C_1 A \in \mathbb{R}^{n\times m}$$

entweder der Nullvektor 0_n oder der kanonische Basisvektor $e_1^{(n)}$ steht.

- Im Fall $a_1 = 0_n \in \mathbb{R}^n$ setzen wir $C_1 = E_n$.

- Im Fall $a_1 \neq 0_n \in \mathbb{R}^n$ gibt es einen kleinsten Zeilenindex $\nu \in 1, \ldots, n$ mit $a_{\nu 1} \neq 0$. Dieser Index sei ν_1. Wir vertauschen die erste und die ν_1-te Zeile von A. Dann normieren wir die erste Zeile. Anschließend annullieren wir die Einträge an den Positionen $(1,2)$ bis $(1,n)$ mit Hilfe der ersten Zeile. Wir setzen
$$C_1 = E^{(n)}_{n;-a_{n1};1} \circ \ldots \circ E^{(n)}_{2;-a_{21};1} \circ E_{a^{-1}_{\nu_1 1};1} \circ E^{(n)}_{1 \leftrightarrow \nu_1} \in \mathrm{GL}(n, \mathbb{R})\,.$$

Der k-te Schritt. Im Fall
$$2 \leq k \leq k_0 = \min\{n, m\}$$
konstruieren wir eine Matrix $C_k \in \mathrm{GL}(n, \mathbb{R})$ derart, dass in der k-ten Spalte von
$$A_k = (a^{(k)}_{\nu\mu}) = (a^{(k)}_1, \ldots, a^{(k-1)}_m) = C_k A_{k-1} \in \mathbb{R}^{n \times m}$$
entweder die k-te Spalte $a^{(k-1)}_k$ von A_{k-1} oder ein kanonischer Basisvektor aus
$$\{e^{(n)}_1, \ldots, e^{(n)}_n\} \setminus \{a^{(k-1)}_1, \ldots, a^{(k-1)}_{k-1}\}$$
steht.

- Wenn $a^{(k-1)}_{\nu k} = 0$ für alle $\nu = k, \ldots, n$ gilt, setzen wir $C_k = E_n$.
- Andernfalls sei
$$\nu_k = \min\{k \leq \nu \leq n \mid a^{(k-1)}_{\nu k} \neq 0\}\,.$$
Wir vertauschen die k-te und die ν_k-te Zeile von A_{k-1}. Dann normieren wir die neue k-te Zeile. Anschließend annullieren wir die Einträge an den Positionen $(k,1)$ bis $(k,k-1)$ und $(k+1,k)$ bis (n,k) mit Hilfe der normierten k-ten Zeile. Wir setzen
$$\begin{aligned} C_k = E^{(n)}_{n;-a^{(k-1)}_{nk};k} \circ \ldots \circ\, & E^{(n)}_{k+1;-a^{(k-1)}_{k+1k};k} \\ & \circ E^{(n)}_{k-1;-a^{(k-1)}_{k-1k};k} \circ \ldots \circ E^{(n)}_{1;-a^{(k-1)}_{1k};k} \\ & \circ E_{1/a^{(k-1)}_{\nu_k k};k} \circ E^{(n)}_{k \leftrightarrow \nu_k}\,. \end{aligned}$$

Der k_0-te Schritt. Wir erhalten $C_1, \ldots, C_{k_0} \in \mathrm{GL}(n, \mathbb{R})$ mit
$$A_{k_0} = C_{k_0} \circ \ldots \circ C_1 \circ A \in \mathrm{T}(\mathbb{R}^{n \times m})\,.$$

Damit ist der Beweis von Aussage (2) beendet. □

Definition 6.7.

(1) *Jede Folge aus endlich vielen elementaren Zeilenumformungen, die aus einer Matrix $A \in \mathbb{R}^{n \times m}$ eine Stufenform aus $\mathrm{U}^{\neg}(\mathbb{R}^{n \times m})$ herstellt, heißt ein* Gauß-Verfahren*.*

(2) *Das im Beweis der Aussage* (1) *des Satzes* 6.6 *beschriebene Verfahren zur Herstellung einer oberen Stufenform heißt das* strikte Gauß-Verfahren*.*

(3) *Jede Folge aus endlich vielen elementaren Zeilenumformungen, die aus einer Matrix* $A \in \mathbb{R}^{n\times m}$ *eine Treppennormalform aus* $\mathrm{T}(\mathbb{R}^{n\times m})$ *herstellt, heißt ein* Gauß-Jordan-Verfahren*.*

(4) *Das im Beweis der Aussage* (2) *des Satzes* 6.6 *beschriebene Verfahren zur Herstellung einer Treppennormalform heißt das* strikte Gauß-Jordan-Verfahren*.*

Manchmal ist es günstiger von den strikten Verfahren abzuweichen und andere elementare Zeilenumformungen einzuschalten.

Wir wenden uns der Untersuchung von Kernen zu. Die Aussage (4) des folgenden Satzes 6.8 ist eine Umformulierung des Satzes 2.16.

Satz 6.8. *Seien* n*,* $m \in \mathbb{N}$*. Sei* $A \in \mathbb{R}^{n\times m}$*. Dann gelten die folgenden Aussagen:*

(1) *Sei* $n = m$*. Wenn* A *invertierbar ist, gilt* $\ker(A) = \{0\}$*.*

(2) $(\forall C \in \mathrm{GL}(n,\mathbb{R}))$: $\ker(CA) = \ker(A)$.

(3) *Die Matrix* $A' \in \mathbb{R}^{n\times m}$ *gehe aus* A *durch elementare Zeilenumformungen hervor. Dann gilt*

$$\ker(A') = \ker(A)\,.$$

(4) *Für* $n < m$ *gilt* $\ker(A) \neq \{0\}$*.*

(5) $(\forall C \in \mathrm{GL}(m,\mathbb{R}))$: $\mathrm{im}(A) = \mathrm{im}(AC)$.

Beweis. Nachweis von (1). Sei $n = m$ und $A \in \mathbb{R}^{n\times n}$ invertierbar. Trivialerweise gilt $\{0\} \subseteq \ker(A)$. Für alle $x \in \ker(A)$ gilt

$$x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}0 = 0\,.$$

Also gilt $\ker(A) = \{0\}$.

Nachweis von (2). Sei $C \in \mathrm{GL}(n,\mathbb{R})$ und $A \in \mathbb{R}^{n\times m}$. Dann gilt

$$\begin{aligned} x \in \ker(A) \quad &\iff \quad Ax = 0 \\ &\iff \quad C(Ax) = 0 \\ &\iff \quad (CA)x = 0 \\ &\iff \quad x \in \ker(CA) \end{aligned}$$

für alle $x \in \mathbb{R}^m$. Damit ist (2) bewiesen.

Aussage (3) folgt aus (2), weil eine elementare Zeilenumformung durch eine linksseitige Multiplikationen mit einer Elementarmatrix beschrieben werden kann.

Aussage (4) ist eine Umformulierung des Satzes 2.16.

Nachweis von (5). Sei $y \in \operatorname{im}(A)$. Dann gibt es $x \in \mathbb{R}^m$ mit $y = Ax$. Nach Voraussetzung gilt $C \in \operatorname{GL}(m, \mathbb{R})$. Es folgt

$$y = Ax = (A \circ (CC^{-1}))x = (AC)(C^{-1}x)\,.$$

Also gilt $y \in \operatorname{im}(AC)$. Wir setzen umgekehrt $y \in \operatorname{im}(AC)$ voraus. Dann gibt es $z \in \mathbb{R}^m$ mit $y = (AC)z = A(Cz)$. Also folgt $y \in \operatorname{im}(A)$. $\square$

Definition 6.9. *Seien n, $m \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times m}$. Dann definieren wir:*

(1) $\operatorname{rg}(A) = \dim(\operatorname{im}(A))$.

(2) $\operatorname{def}(A) = \dim(\ker(A))$.

Wir nennen $\operatorname{rg}(A)$ *den* Rang *und* $\operatorname{def}(A)$ *den* Defekt *der Matrix A.*

Satz 6.10 (Rang, Defekt, Dimensionsformel)**.** *Seien n, $m \in \mathbb{N}$. Dann gelten die folgenden Aussagen* (1) *bis* (7)*.*

(1) $(\forall T \in \mathrm{T}(\mathbb{R}^{n\times m})): \quad \operatorname{rg}(T^t) = \operatorname{rg}(T) = \#(PV(T))$.

(2) $(\forall C \in \operatorname{GL}(n, \mathbb{R}))(\forall A \in \mathbb{R}^{n\times m}): \quad \operatorname{rg}(CA) = \operatorname{rg}(A)$.

(3) $(\forall A \in \mathbb{R}^{n\times m}): \quad \operatorname{rg}(A^t) = \operatorname{rg}(A)$.

(4) $(\forall T \in \mathrm{T}(\mathbb{R}^{n\times m})): \quad \operatorname{def}(T) = m - \#(PV(T))$.

(5) $(\forall C \in \operatorname{GL}(n, \mathbb{R}))(\forall A \in \mathbb{R}^{n\times m}): \quad \operatorname{def}(CA) = \operatorname{def}(A)$.

(6) $(\forall A \in \mathbb{R}^{n\times m}): \quad 0 \leq \operatorname{rg}(A) \leq \min\{m, n\}$.

(7) $(\forall A \in \mathbb{R}^{n\times m}): \quad \operatorname{rg}(A) + \operatorname{def}(A) = m$.

Die Gleichung in Aussage (7) *ist die* Dimensionsformel.

Beweis. Nachweis von (1). Sei $T \in \mathrm{T}(\mathbb{R}^{n\times m})$ eine Treppennormalform mit $r = \#(PV(T)) \in \mathbb{N}$. Dann wird $\operatorname{im}(T) \subseteq \mathbb{R}^n$ von den ersten r Vektoren $e_1^{(n)}, \ldots, e_r^{(n)}$ der kanonischen Basis $\mathcal{E}_n$ aufgespannt. Es gilt

$$\operatorname{rg}(T) = \dim(\operatorname{im}(T)) = r = \#(PV(T))\,.$$

Seien $z_1, \ldots, z_n \in \mathbb{R}_m$ die Zeilen von T. Wir transponieren T. Dann bilden die ersten r Spalten $z_1^t, \ldots, z_r^t \in \mathbb{R}^m$ von T^t eine Basis von $\operatorname{im}(T^t) \subseteq \mathbb{R}^m$. Also gilt

$$\operatorname{rg}(T^t) = \dim(\operatorname{im}(T^t)) = r = \operatorname{rg}(T)\,.$$

Im Fall $\operatorname{rg}(T) = 0$ ist Aussage (1) trivialerweise erfüllt.

Nachweis von (2). Offenbar gilt $\operatorname{rg}(A) = 0$ genau dann, wenn $A = 0_{n\times m}$ gilt. Daher ist (2) im Fall $\operatorname{rg}(A) = 0$ klar. Wir setzen nun $\operatorname{rg}(A) = r \in \mathbb{N}$ voraus.

Dann gibt es eine Basis $\{s_1, \ldots, s_r\}$ von $\mathrm{im}(A)$. Wegen $C \in \mathrm{GL}(n, \mathbb{R})$ ist die Menge $\{Cs_1, \ldots, Cs_r\}$ linear unabhängig. Also gilt

$$\mathrm{rg}(A) = r \leq \mathrm{rg}(CA)\,.$$

Analog folgt für CA und C^{-1} anstelle von A respektive C, dass

$$\mathrm{rg}(CA) \leq \mathrm{rg}(C^{-1}(CA)) = \mathrm{rg}(A)$$

gilt. Also folgt $\mathrm{rg}(CA) = \mathrm{rg}(A)$.

Nachweis von (3). Nach Satz 6.6 gibt es $C \in \mathrm{GL}(n, \mathbb{R})$ und $T \in \mathrm{T}(\mathbb{R}^{n\times m})$ mit $A = CT$. Folglich gilt $A^t = T^t C^t$. Mit C ist auch C^t invertierbar. Nach Satz 6.8, Aussage (5) gilt

$$\mathrm{im}(A^t) = \mathrm{im}(T^t C^t) = \mathrm{im}(T^t)\,.$$

Mit den bereits bewiesenen Aussagen (1) und (2) erhalten wir

$$\mathrm{rg}(A^t) = \mathrm{rg}(T^t) = \mathrm{rg}(T) = \mathrm{rg}(CT) = \mathrm{rg}(A)\,.$$

Nachweis von (4) *bis* (6). Rückwärtselimination liefert die Gültigkeit von (4). Aussage (5) ist eine Umformulierung von Satz 6.8, (2). Aussage (6) folgt aus (3).

Nachweis von (7). Die Dimensionsformel gilt nach (4) für alle Matrizen aus $\mathrm{T}(\mathbb{R}^{n\times m})$. Für eine beliebige Matrix $A \in \mathbb{R}^{n\times m}$ gibt es nach Satz 6.6 Matrizen $C \in \mathrm{GL}(n, \mathbb{R})$ und $T \in \mathrm{T}(\mathbb{R}^{n\times m})$ mit $A = CT$. Nach den Aussagen (2) und (5) gilt die Dimensionsformel daher auch für A. □

Satz 6.11 (Streichen und Ergänzen)**.** *Seien $m, n \in \mathbb{N}$ und $T \in T(\mathbb{R}^{n\times m})$. Eine Basis von $\ker(T) \subseteq \mathbb{R}^m$ kann auf folgende Weise berechnet werden:*

(1) *Sei $r = \mathrm{rg}(T) = \#(PV(T))$.*

(2) *Sei $T' \in \mathrm{T}(\mathbb{R}^{r\times m})$ die Matrix, die durch Streichen aller Nullzeilen aus T hervorgeht.*

(3) *Die Matrix*

$$T'' = (t''_{\mu\lambda})_{\substack{\mu=1,\ldots,m\\ \lambda=1,\ldots,m}} \in \mathbb{R}^{m\times m}$$

entstehe aus T', indem $(m - r)$ Zeilen aus $\mathbb{R}_m$, die an genau einer Stelle eine -1 und sonst lauter Nullen enthalten, derart ergänzt werden, dass

$$t''_{\mu\mu} = \pm 1$$

für alle $\mu = 1, \ldots m$ gilt.

(4) *Die $(m - r)$ Spalten von T'', die eine ergänzte -1 enthalten, bilden eine Basis $\mathcal{A}(\ker(T))$ von $\ker(T)$.*

(5) *Nach der Dimensionsformel gilt*

$$\dim(\ker(T)) = m - r\,.$$

Siehe Satz 6.10*, Aussage* (7)*.*

Wir sagen, dass die Matrix $T'' \in \mathbb{R}^{m\times m}$ und die Basis $\mathcal{A}(\ker(T)) \subseteq \mathbb{R}^m$ durch Streichen und Ergänzen aus T hervorgehen*.*

Beweis. Wir beweisen Aussage (4). Sei $T = (t_{\nu\mu}) \in \mathrm{T}(\mathbb{R}^{n\times m})$. In den Fällen $r = 0$ und $m - r = 0$ ist Aussage (4) klar. Wir betrachten den Fall $1 \leq r = \mathrm{rg}(T) \leq \min\{n, m\}$ und $1 \leq k = \mathrm{def}(T) = m - r$. Seien $1 \leq i_1 < \ldots < i_r \leq m$ die Indizes der Pivotspalten und $1 \leq j_1 < \ldots < j_k \leq m$ die Indizes der übrigen Spalten. Die quadratische Matrix

$$T'' = (t''_{\mu\lambda}) = (t''_1, \ldots, t''_m) \in \mathrm{T}(\mathbb{R}^{m\times m})$$

mit den Spalten $t''_1, \ldots, t''_m \in \mathbb{R}^m$ entstehe aus T durch Streichen und Ergänzen. Siehe (2) und (3). Wir zeigen nun, dass die k linear unabhängigen Vektoren $t''_{j_1}, \ldots, t''_{j_k} \in \mathbb{R}^m$ im Kern der Matrix T enthalten sind.

(i) Für $\nu = 1, \ldots, r$ gilt

$$t_{\nu i_\rho} = \delta_{\nu\rho}\,, \quad \rho = 1, \ldots, r\,.$$

(ii) Für $\mu = 1, \ldots, m$ und $\sigma = 1, \ldots, k$ gilt

$$t''_{\mu j_\sigma} = \begin{cases} -\delta_{\kappa\sigma}\,, & \mu = j_\kappa\,, \quad \kappa = 1, \ldots, k\,, \\ \ \ t_{\rho j_\sigma}\,, & \mu = i_\rho\,, \quad \rho = 1, \ldots, r\,. \end{cases}$$

Für $\nu = 1, \ldots, r$ und $\sigma = 1, \ldots, k$ erhalten wir

$$\begin{aligned} \sum_{\mu=1}^{m} t_{\nu\mu} t''_{\mu j_\sigma} &= \sum_{\rho=1}^{r} t_{\nu i_\rho} t''_{i_\rho j_\sigma} + \sum_{\kappa=1}^{k} t_{\nu j_\kappa} t''_{j_\kappa j_\sigma} \\ &= \sum_{\rho=1}^{r} \delta_{\nu\rho} \cdot t_{\rho j_\sigma} + \sum_{\kappa=1}^{k} t_{\nu j_\kappa} \cdot (-\delta_{\kappa\sigma}) \\ &= t_{\nu j_\sigma} + (-t_{\nu j_\sigma}) = 0\,. \end{aligned}$$

Für $\nu = r + 1, \ldots, m$ und $\sigma = 1, \ldots, k$ erhalten wir

$$\sum_{\mu=1}^{m} t_{\nu\mu} t''_{\mu j_\sigma} = \sum_{\mu=1}^{m} 0 \cdot t''_{\mu j_\sigma} = 0\,.$$

Also ist $\{t''_{j_1}, \ldots, t''_{j_k}\}$ eine Basis von $\ker(T)$. □

Beispiel 6.12. Gegeben sei die Treppennormalform

$$T = (t_1, t_2, t_3, t_4, t_5) = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in T(\mathbb{R}^{4\times 5})\,.$$

Die Vektoren t_1, t_3, $t_5 \in \mathbb{R}^4$ sind die Pivotspalten von T. Es gilt $\mathrm{rg}(T) = 3$. Die Vektoren

$$z_1 = t_2'' = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^5\,, \quad z_2 = t_4'' = \begin{pmatrix} 4 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \in \mathbb{R}^5$$

bilden nach Satz 6.11 eine Basis von $\ker(T) \subseteq \mathbb{R}^5$. Siehe Beispiel 6.21. □

Nach diesen Vorbereitungen können wir den angekündigten Zerlegungssatz für Matrizen beweisen. Aus der Einzigkeitsaussage des Zerlegungssatzes 6.13 folgt insbesondere, dass eine Treppennormalform nicht von den elementaren Zeilenumformungen abhängt. Die Einzigkeit der Treppennormalform einer Matrix gestattet es, die Pivotspalten einer beliebigen Matrix zu definieren.

Satz und Definition 6.13 (Gauß-Jordan-Zerlegung)**.** *Seien $m, n \in \mathbb{N}$ und*

$$A = (a_1, \ldots, a_m) \in \mathbb{R}^{n\times m}$$

mit den Spalten $a_1, \ldots, a_m \in \mathbb{R}^n$ beliebig gegeben.

(1) *Es gibt Matrizen $C \in \mathrm{GL}(n, \mathbb{R})$ und $T \in \mathrm{T}(\mathbb{R}^{n\times m})$ mit*

$$A = C \circ T\,.$$

Eine solche Zerlegung heißt eine Gauß-Jordan-Zerlegung *der Matrix A. Die Treppennormalform T ist durch A eindeutig bestimmt. Sie heißt die* Treppennormalform *von A und wird mit $T(A)$ bezeichnet.*

(2) *Die Matrix $T(A)$ ist die einzige Treppennormalform in $T(\mathbb{R}^{n\times m})$, die aus A durch elementare Zeilenumformungen hervorgeht.*

(3) *Sei $r = \mathrm{rg}(A) \in \mathbb{N}$. Seien $t_{\mu_1}, \ldots, t_{\mu_r}$ mit $1 \leq \mu_1 < \ldots < \mu_r \leq m$ die r Pivotspalten von*

$$T(A) = (t_1, \ldots, t_m)\,.$$

Dann heißen die r Spalten $a_{\mu_1}, \ldots, a_{\mu_r}$ die Pivotspalten *der Matrix A. Im Fall* $\mathrm{rg}(A) = 0$ *besitzt A keine Pivotspalten.*

(4) *Die invertierbare Matrix C ist im Allgemeinen* nicht *durch A eindeutig bestimmt, wie das Beispiel*

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \alpha \in \mathbb{R}, \quad \beta \in \mathbb{R}^{\times}$$

zeigt. Im Spezialfall $\mathrm{rg}(A) = n$ *ist die invertierbare Matrix C eindeutig durch die Matrix A bestimmt.*

(5) *Werden die Zeilenumformungen, die A in $T(A)$ überführen, in derselben Reihenfolge auf die Einheitsmatrix E_n angewendet, so ergibt sich aus E_n die inverse Matrix C^{-1} einer invertierbaren Matrix C mit $A = C \circ T(A)$. Schematisch:*

$$\begin{array}{c|c} A & E_n \\ \hline T(A) & C^{-1} \end{array}$$

$$\begin{array}{c|c|c} C^{-1} & E_n & T(A) \\ \hline E_n & C & A \end{array}$$

Die Matrix C hängt von den elementaren Zeilenumformungen ab, die A in $T(A)$ überführen. Siehe Beispiel 6.21.

Beweis. Nachweis von (1). Sei $A \in \mathbb{R}^{n\times m}$ beliebig gegeben. Nach Satz 6.6 gibt es $C \in \mathrm{GL}(n, \mathbb{R})$ und $T \in \mathrm{T}(\mathbb{R}^{n\times m})$ mit

$$A = CT\,.$$

Nachweis von (2). Wir zeigen, dass die Treppennormalform T durch A eindeutig bestimmt ist. Seien C_1, $C_2 \in \mathrm{GL}(n, \mathbb{R})$ und T_1, $T_2 \in \mathrm{T}(\mathbb{R}^{n\times m})$ mit

$$A = C_1T_1 = C_2T_2$$

gegeben. Nach Satz 6.8 gilt

$$\ker(A) = \ker(T_1) = \ker(T_2)\,.$$

Eine Basis des Kernes einer Treppennormalform ergibt sich durch Streichen und Ergänzen. Dies zeigt, dass die Spalten von T_1 und T_2 übereinstimmen, die keine Pivotspalten sind. Nach Satz 6.10 gilt

$$\mathrm{rg}(A) = \#(PV(T_1)) = \#(PV(T_2))\,.$$

Die Pivotspalten von T_1 und T_2 sind kanonische Basisvektoren. Also stimmen auch die Pivotspalten von T_1 und T_2 überein.

Nachweis von (4). Sei $A = (a_1, \ldots, a_m) \in \mathbb{R}^{n\times m}$ mit $\mathrm{rg}(A) = n$ gegeben. Nach (1) gibt es $C = (c_1, \ldots, c_n) \in \mathrm{GL}(n, \mathbb{R})$ mit

$$A = C \circ T(A)\,.$$

Für die n Pivotspalten a_{μ_ν} von A und die n Pivotspalten $t_{\mu_\nu} = e_\nu^{(n)}$ von $T(A)$ gilt

$$a_{\mu_\nu} = Ct_{\mu_\nu} = Ce_\nu^{(n)} = c_\nu \,.$$

Daher wird C im Fall $\mathrm{rg}(A) = n$ eindeutig durch A bestimmt.

Aussage (5) folgt aus dem Verfahren von Gauß-Jordan. □

Wir charakterisieren die invertierbaren Matrizen. Das Verfahren von Gauß-Jordan ergibt ein Verfahren zur Berechnung der inversen Matrix.

Satz 6.14 (Charakterisierung invertierbarer Matrizen)**.** *Sei $n \in \mathbb{N}$. Sei $A = (a_1, \dots, a_n) \in \mathbb{R}^{n\times n}$ eine $(n \times n)$-Matrix mit den Spalten $a_1, \dots, a_n \in \mathbb{R}^n$. Dann sind die folgenden Aussagen* (1) *bis* (7) *äquivalent.*

(1) $\ker(A) = \{0\}$.

(2) $\mathrm{im}(A) = \mathbb{R}^n$.

(3) $\mathcal{A} = \{a_1, \dots, a_n\}$ *ist eine Basis des* $\mathbb{R}^n$.

(4) $(\exists B \in \mathbb{R}^{n\times n}): \quad AB = E_n$.

(5) $(\exists B \in \mathbb{R}^{n\times n}): \quad BA = E_n$.

(6) $T(A) = E_n$.

(7) $A \in \mathrm{GL}(n, \mathbb{R})$.

Die Matrizen $B \in \mathbb{R}^n$ in den Aussagen (4) *und* (5) *sind eindeutig durch A bestimmt. Es gilt*

$$B = A^{-1} \,.$$

Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann invertierbar, wenn A durch elementare Zeilenumformungen in die Einheitsmatrix $E_n \in \mathbb{R}^{n\times n}$ überführt werden kann. Werden diese elementaren Zeilenumformungen in derselben Reihenfolge auf die Matrix E_n angewendet, so entsteht die inversen Matrix A^{-1}.

$$\begin{array}{c|c} A & E_n \\ \hline E_n & A^{-1} \end{array}$$

Beweis. Nach der Dimensionsformel aus Satz 6.10 sind (1) und (2) äquivalent. Aussage (1) bedeutet, dass $a_1, \dots, a_n$ linear unabhängig sind. Aussage (2) bedeutet $[\mathcal{A}] = \mathbb{R}^n$. Also sind (1), (2) und (3) äquivalent.

Nach dem Zerlegungssatz 6.13 sind (6) und (7) äquivalent. Aus (7) folgt (1).

Aus (2) folgt (4). Wir setzen (4) voraus. Aus $AB = E_n$ folgt $\ker(B) = \{0\}$. Folglich gibt es $C \in \mathbb{R}^{n\times n}$ mit $BC = E_n$. Wir erhalten

$$BA = (BA)(BC) = B(AB)C = BC = E_n \,.$$

Also folgen (5) und (7) aus (4). Bisher haben wir die Äquivalenz von (1), (2), (3), (4), (6) und (7) bewiesen.

Wir setzen (5) voraus. Aus $BA = E_n$ folgt $A^t B^t = E_n^t = E_n$. Nach der bereits bewiesenen Äquivalenz von (4) und (7) ist die Matrix A^t invertierbar. Damit erhalten wir $(A^t)^{-1} = B^t$. Also gilt $B^t A^t = E_n$. Es folgt $AB = E_n$. Also folgt (4) aus (5). Damit haben wir Äquivalenz von (1) bis (7) bewiesen. □

Das Gauß-Jordan-Verfahren liefert nach 6.13 und 6.14 eine Zerlegung einer invertierbaren Matrix in ein Produkt aus endlich vielen Elementarmatrizen. Allerdings ist eine solche Zerlegung nicht eindeutig bestimmt.

Satz 6.15 (Zerlegung in Elementarmatrizen)**.** *Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann invertierbar, wenn es Elementarmatrizen $C_1, \ldots, C_k \in \mathrm{Z}(n, \mathbb{R})$ mit*

$$A = C_1 \cdots C_k$$

gibt. Eine solche Zerlegung invertierbarer Matrizen ist nicht eindeutig bestimmt.

Beispiel 6.16. Die Matrix $A \in \mathbb{R}^{2\times 2}$ mit

$$A = \begin{pmatrix} \frac{1}{5} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

ist invertierbar mit

$$A^{-1} = \begin{pmatrix} 180 & -210 \\ -210 & 252 \end{pmatrix} .$$

Das Gauß-Verfahren liefert

$\frac{1}{5}$	$\frac{1}{6}$	1	0	
$\frac{1}{6}$	$\frac{1}{7}$	0	1	
1	$\frac{5}{6}$	5	0	$5\,Z_1$
$\frac{1}{6}$	$\frac{1}{7}$	0	1	
1	$\frac{5}{6}$	5	0	
0	$\frac{1}{252}$	$-\frac{5}{6}$	1	$Z_2 - \frac{1}{6}Z_1$
1	$\frac{5}{6}$	5	0	
0	1	-210	252	$252\,Z_2$
1	0	180	-210	$Z_1 - \frac{5}{6}Z_2$
0	1	-210	252	

In Matrixschreibweise erhalten wir

$$\begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{5} & \frac{1}{6} & 1 & 0 & \frac{7}{6} \\ \frac{1}{6} & \frac{1}{7} & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{6} & 5 & 0 & \frac{35}{6} \\ \frac{1}{6} & \frac{1}{7} & 0 & 1 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ -\frac{1}{6} & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} & 5 & 0 & \frac{35}{6} \\ \frac{1}{6} & \frac{1}{7} & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{6} & 5 & 0 & \frac{35}{6} \\ 0 & \frac{1}{252} & -\frac{5}{6} & 1 & \frac{1}{36} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 252 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} & 5 & 0 & \frac{35}{6} \\ 0 & \frac{1}{252} & -\frac{5}{6} & 1 & \frac{1}{36} \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{6} & 5 & 0 & \frac{35}{6} \\ 0 & 1 & -210 & 252 & 7 \end{pmatrix},$$

$$\begin{pmatrix} 1 & -\frac{5}{6} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} & 5 & 0 & \frac{35}{6} \\ 0 & 1 & -210 & 252 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 180 & -210 & 0 \\ 0 & 1 & -210 & 252 & 7 \end{pmatrix}.$$

Dementsprechend erhalten die Zerlegung

$$\begin{pmatrix} \frac{1}{5} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{7} \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{6} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{252} \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} \\ 0 & 1 \end{pmatrix}.$$

Solche eine Zerlegung hängt offenbar von den elementaren Zeilenumformungen ab, die zur Herstellung der Treppennormalform verwendet worden sind. □

Nach diesen Vorbereitungen können wir die bereits erwähnten Sätze zur Bestimmung von Basen des Kernes und des Bildes einer Matrix formulieren.

Satz 6.17 (Berechnung einer Basis des Kernes)**.** *Seien $m, n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times m}$. Eine Basis von $\ker(A)$ kann auf folgende Weise berechnet werden:*

(1) *Berechnung des Ranges $r = \mathrm{rg}(A)$ und der Treppennormalform $T(A)$ der Matrix A mit dem Verfahren von Gauß-Jordan.*

(2) *Sei $T'(A) \in \mathrm{T}(\mathbb{R}^{r\times m})$ die Matrix, die durch Streichen aller Nullzeilen aus $T(A)$ hervorgeht.*

(3) *Die Matrix*

$$T''(A) = (t''_{\nu\mu})_{\substack{\nu=1,\ldots,m \\ \mu=1,\ldots,m}} \in \mathbb{R}^{m\times m}$$

entstehe aus $T'(A)$, indem $(m-r)$ Zeilen aus $\mathbb{R}_m$, die an genau einer Stelle eine -1 und sonst lauter Nullen enthalten, derart ergänzt werden, dass $t''_{\nu\nu} = \pm 1$ für alle $\nu = 1, \ldots m$ gilt.

(4) *Die Spalten von $T''(A)$, die eine ergänzte -1 enthalten, bilden eine Basis von $\ker(A)$.*

Satz 6.18 (Extraktion einer Basis)**.** *Seien* $m, n \in \mathbb{N}$ *und* $a_1, \ldots, a_m \in \mathbb{R}^n$ *gegeben. Sei*

$$X = [a_1, \ldots, a_m] = [\{a_1, \ldots, a_m\}] \subseteq \mathbb{R}^n$$

der Teilvektorraum, der von den Vektoren $a_1, \ldots, a_m$ *erzeugt wird. Sei weiter*

$$A = (a_1, \ldots, a_m) \in \mathbb{R}^{n \times m}$$

die $(n \times m)$*-Matrix mit den Spalten* $a_1, \ldots, a_m$*. Eine Basis von* $X = \mathrm{im}(A)$ *kann aus den Vektoren* $a_1, \ldots, a_m$ *auf folgende Weise extrahiert werden:*

(1) *Berechnung des Ranges* $r = \mathrm{rg}(A)$ *und der Treppennormalform* $T(A)$ *der Matrix* A *mit dem Verfahren von Gauß-Jordan.*

(2) $r = \mathrm{rg}(A) = \mathrm{rg}(T(A)) = \#(PV(T(A)))$.

(3) *Seien* $t_{\mu_1}, \ldots, t_{\mu_r}$ *die Pivotspalten der Treppenormalform von* $T(A)$*. Dann heißen die Spaltenvektoren* $a_{\mu_1}, \ldots, a_{\mu_r}$ *die* Pivotspalten *der Matrix* A*.*

(4) *Die Pivotspalten* $a_{\mu_1}, \ldots, a_{\mu_r}$ *der Matrix* A *bilden eine Basis*

$$\mathcal{A} = \{a_{\mu_1}, \ldots, a_{\mu_r}\}$$

des Teilraumes $X = \mathrm{im}(A)$*.*

Wir verwenden vorstehende Ergebnisse, um ein Lösungsverfahren für lineare inhomogene Gleichungsysteme zu formulieren.

Satz 6.19 (Lösbarkeit linearer Gleichungssysteme)**.** *Seien* $n,\ m \in \mathbb{N}$*. Seien* $A = (a_1, \ldots, a_m) \in \mathbb{R}^{n \times m}$ *und* $b \in \mathbb{R}^m$ *gegeben. Dann sind die folgenden Aussagen* (1)*,* (2)*,* (3) *äqivalent.*

(1) $Ax = b$ *ist lösbar.*

(2) $b \in \mathrm{im}(A)$.

(3) $\mathrm{rg}(A) = \mathrm{rg}(A')$.

Dabei ist

$$A' = (A, b) = (a_1, \ldots, a_m, b) \in \mathbb{R}^{n \times (m+1)}$$

die $n \times (m+1)$*-Matrix, die aus* A *durch Hinzunahme von* b *als einer* $(m+1)$*-ten Spalte hervorgeht. Die Matrix* (A, b) *heißt dann auch die* erweiterte Matrix*. Rang und Treppennormalform der erweiterten Matrix* (A, b) *bezeichnen wir mit* $\mathrm{rg}(A, b)$ *respektive* $T(A, b)$*. Offenbar sind homogene lineare Gleichungssystem* $Ax = 0$ *stets lösbar.*

Satz 6.20. *Seien n, $m \in \mathbb{N}$. Seien $A = (a_1, \dots, a_m) \in \mathbb{R}^{n\times m}$ und $b \in \operatorname{im}(A)$ gegeben. Sei $L \subseteq \mathbb{R}^m$ die Menge aller Lösungen von $Ax = b$. Dann gelten:*

(1) *$Ax = b$ ist lösbar.*

(2) *Für alle $\xi \in L$ gilt*

$$L = \xi + \ker(A) = \{\, x \in \mathbb{R}^m \mid (\exists z \in \ker(A)) : \ x = \xi + z \,\} .$$

(3) *Für Aussage (2) ist folgende Ausdrucksweise üblich. Ein fest gewählte $\xi \in L$ heißt eine* spezielle Lösung. *Jede Lösung des inhomogenen Gleichungssystems ist die Summe einer speziellen Lösung und einer Lösung des homogenen Gleichungssystems.*

(4) *Sei $r = \operatorname{rg}(A)$ und $T'(A, b) \in \mathrm{T}(\mathbb{R}^{r\times(m+1)})$ die Matrix, die durch Streichen aller Nullzeilen aus $T(A, b)$ hervorgeht.*

(5) *Die Matrix*

$$T''(A, b) = (t''_{\nu\mu})_{\substack{\nu=1,\dots,m\\ \mu=1,\dots,m+1}} \in \mathbb{R}^{m\times(m+1)}$$

entstehe aus $T'(A, b)$, indem $(m-r)$ Zeilen aus $\mathbb{R}^{m+1}$, die an genau einer Stelle eine -1 und sonst lauter Nullen enthalten, derart ergänzt werden, dass $t''_{\nu\nu} = \pm 1$ für alle $\nu = 1, \dots, m$ gilt.

(6) *Die Spalten von $T''(A, b)$, die eine ergänzte -1 enthalten, bilden eine Basis von* $\ker(A)$.

(7) *Die $(m + 1)$-te Spalte von $T''(A, b)$ ist ein Element von L.*

Beispiel 6.21. Wir betrachten die Matrix $A \in \mathbb{R}^{4\times 5}$ und den Vektor $b \in \mathbb{R}^4$ mit

$$A = (a_1, a_2, a_3, a_4, a_5) = \begin{pmatrix} 1 & 2 & 2 & 4 & 3 \\ 2 & 4 & 3 & 8 & 4 \\ 3 & 6 & 4 & 12 & 5 \\ 4 & 8 & 5 & 16 & 4 \end{pmatrix} , \quad b = \begin{pmatrix} 7 \\ 10 \\ 13 \\ 12 \end{pmatrix} .$$

Die Zerlegungen

$$\begin{pmatrix} 1 & 2 & 2 & 4 & 3 \\ 2 & 4 & 3 & 8 & 4 \\ 3 & 6 & 4 & 12 & 5 \\ 4 & 8 & 5 & 16 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 0 \\ 3 & 4 & 5 & \alpha \\ 4 & 5 & 4 & \beta \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

mit $\alpha \in \mathbb{R}^\times$ und $\beta \in \mathbb{R}$ sind Gauß-Jordan-Zerlegungen von A. Es gibt auch noch andere Gauß-Jordan-Zerlegungen von A. Den Rang, eine Basis des Kernes und die Pivotspalten von A können wir an der Treppennormalform

$$T(A) = (t_1, t_2, t_3, t_4, t_5) = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{T}(\mathbb{R}^{4\times 5})$$

ablesen. Es gilt $\mathrm{rg}(A) = \mathrm{rg}(T(A)) = 3$. Streichen und Ergänzen liefert

$$T''(A) = (t''_1, t''_2, t''_3, t''_4, t''_5) = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{5\times 5}.$$

Die Vektoren

$$z_1 = t''_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^5, \quad z_2 = t''_4 = \begin{pmatrix} 4 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \in \mathbb{R}^5$$

bilden nach Satz 6.17 eine Basis von $\ker(A) = \ker(T(A))$. Siehe Beispiel 6.12. Die Vektoren t_1, t_3, t_5 sind die Pivotspalten von $T(A)$. Daher sind

$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \in \mathbb{R}^4, \quad a_3 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \in \mathbb{R}^4, \quad a_5 = \begin{pmatrix} 3 \\ 4 \\ 5 \\ 4 \end{pmatrix} \in \mathbb{R}^4$$

die Pivotspalten der gegebenen Matrix A. Nach Satz 6.18 bilden die Pivotspalten von A eine Basis $\mathcal{A} = \{a_1, a_3, a_5\}$ von $\mathrm{im}(A)$. Aus der Treppennormalform

$$T(A, b) = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 6}.$$

der erweiterten Matrix $(A, b) = (a_1, a_2, a_3, a_4, a_5, b)$ folgt, dass das inhomogene lineare Gleichungssystem $Ax = b$ lösbar ist, denn es gilt

$$\mathrm{rg}(A, b) = \mathrm{rg}(T(A, b)) = \mathrm{rg}(T(A)) = \mathrm{rg}(A).$$

Siehe Satz 6.19. Streichen und Ergänzen liefert

$$T''(A, b) = (t''_1, t''_2, t''_3, t''_4, t''_5, t''_6) = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{5\times 6}.$$

Im Falle der Lösbarkeit gilt $T''(A,b) = (T''(A), t''_6)$, wobei

$$\xi = t''_6 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \in \mathbb{R}^5$$

nach Satz 6.20 eine spezielle Lösung von $Ax = b$ ist. Demnach bilden alle Vektoren $x \in \mathbb{R}^5$ der Form

$$x = \xi + \lambda_1 z_1 + \lambda_2 z_2 \,, \quad \lambda_1 \in \mathbb{R} \,, \quad \lambda_2 \in \mathbb{R}$$

die Lösungsmenge von $Ax = b$. $\square$

Schema 6.22.

- $A = (a_1, \dots, a_m) \in \mathbb{R}^{n \times m}$ mit $r = \operatorname{rg}(A) \in \mathbb{N}$
- $t_{i_1}, \dots, t_{i_r}$ mit $1 \leq i_1 < \dots < i_r \leq m$ seien die Pivotspalten von $T(A)$
- $\mathcal{A} = \{a_{i_1}, \dots, a_{i_r}\}$ ist eine Basis von $\operatorname{im}(A)$
- $(a_{i_1}, \dots, a_{i_r}) \in \mathbb{R}^{n \times r}$
- $b \in \mathbb{R}^n$
- $L = \{\, x \in \mathbb{R}^m \mid Ax = b \,\}$

Kriterium: $L \neq \emptyset \iff b \in \operatorname{im}(A) \iff \operatorname{rg}(T(A,b)) = \operatorname{rg}(T(A))$

- $b \in \operatorname{im}(A) \implies \begin{cases} T'(A,b) = (T'(A), b_{\mathcal{A}}) \in \mathbb{R}^{r \times (m+1)}, \\ (a_{i_1}, \dots, a_{i_r})\, b_{\mathcal{A}} = b \end{cases}$
- $b \in \operatorname{im}(A) \implies \begin{cases} T''(A,b) = (T''(A), \xi) \in \mathbb{R}^{m \times (m+1)}, \\ A\xi = b \end{cases}$
- $b \in \operatorname{im}(A) \implies L = \xi + \ker(A) = \{\, \xi + \eta \mid \eta \in \ker(A) \,\}$

- $\operatorname{def}(A) = k = m - r \in \mathbb{N}$
- $t_{j_1}, \dots, t_{j_k}$ mit $1 \leq j_1 < \dots < j_k \leq m$ seien die übrigen Spalten von $T(A)$
- $T''(A) = (t''_1, \dots, t''_m) \in \mathbb{R}^{m \times m}$
- $\mathcal{N} = \{t''_{j_1}, \dots, t''_{j_k}\}$ ist eine Basis von $\ker(A) = \ker(T(A))$

7 Koordinaten und Matrixdarstellungen

Im Folgenden seien $X \subseteq \mathbb{R}^{m'}$, $Y \subseteq \mathbb{R}^{n'}$ und $Z \subseteq \mathbb{R}^{k'}$ beliebige Teilvektorräume. Wir wollen dies so verstehen, dass Gleichheit $X = Y$ die Gleichheit $m' = n'$ impliziert. Sei X nicht der Nullraum. Dann ist eine lineare Abbildung $\varphi : X \to Y$ durch die Bilder der Vektoren einer Basis von X eindeutig bestimmt. Wenn Y ebenfalls nicht der Nullraum ist, dann kann φ bezüglich einer Basis $\mathcal{A}$ von X und einer Basis $\mathcal{B}$ von Y durch eine Matrix $M_{\mathcal{AB}}$ dargestellt werden. Auf diese Weise lassen sich lineare Abbildungen mit dem Matrixkalkül untersuchen. Wir beginnen mit der Definition der Koordinatenvektoren.

Satz und Definition 7.1. *Sei $m \in \mathbb{N}$ und $\mathcal{E}_m = \{e_1^{(m)}, \dots, e_m^{(m)}\}$ die kanonische Basis des $\mathbb{R}^m$. Sei weiter $\mathcal{A} = \{a_1, \dots, a_m\}$ eine beliebige Basis von X.*

(1) *Jeder Vektor $x \in X$ besitzt eine eindeutig bestimme Darstellung*

$$x = \sum_{\mu=1}^{m} \alpha_\mu a_\mu$$

mit $\alpha_1, \dots, \alpha_m \in \mathbb{R}$. Der Spaltenvektor

$$x_{\mathcal{A}} = \sum_{\mu=1}^{m} \alpha_\mu e_\mu^{(m)} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \in \mathbb{R}^m$$

heißt der Koordinatenvektor *von x bezüglich $\mathcal{A}$. Dementsprechend heißen die Komponenten $\alpha_1, \dots, \alpha_m$ des Spaltenvektors $x_{\mathcal{A}}$ die* Koordinaten *von x bezüglich $\mathcal{A}$.*

(2) *Sei $A = (a_1, \dots, a_m) \in \mathbb{R}^{m' \times m}$ die Matrix mit den Spalten $a_1, \dots, a_m$. Dann gelten* (2.1), (2.2), (2.3), (2.4).

(2.1) $\mathrm{rg}(A) = m \leq m'$.

(2.2) $X = \mathrm{im}(A) = [a_1, \dots, a_m] \subseteq \mathbb{R}^{m'}$.

(2.3) *Für alle $x \in X$ gilt*

$$x = Ax_{\mathcal{A}} = (a_1, \dots a_m) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \sum_{\mu=1}^{m} \alpha_\mu a_\mu \,.$$

(2.4) *Für jeden Vektor $x \in X$ ist der Koordinatenvektor $x_{\mathcal{A}} \in \mathbb{R}^m$ ist die einzige Lösung $\xi \in \mathbb{R}^m$ des inhomogenen linearen Gleichungssystems*

$$A\xi = x \,.$$

(3) *Sei $x \in X$ beliebig gegeben. Der Koordinatenvektor $x_{\mathcal{A}} \in \mathbb{R}^m$ entsteht aus der Treppennormalform $T(A, x) \in \mathrm{T}(\mathbb{R}^{m' \times (m+1)})$ durch Streichen der $(m' - m)$ Nullzeilen und der ersten m Spalten.*

(4) *Es gilt*

$$(a_\mu)_{\mathcal{A}} = e_\mu^{(m)}$$

für alle $\mu = 1, \dots, m$.

Wir heben ausdrücklich hervor, dass der Koordinatenvektor $x_{\mathcal{A}} \in \mathbb{R}^m$ vom Vektor $x \in X \subseteq \mathbb{R}^{m'}$ im Allgemeinen sorgfältig unterschieden werden muss. Im Fall $m < m'$ ist dies offenkundig. Wir betrachten ein Beispiel mit $m = m' = 3$. Dagegen gilt stets $x = x_{\mathcal{E}}$, wenn $\mathcal{E}$ die kanonische Basis ist.

Beispiel 7.2. Die Vektoren

$$a_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad a_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

bilden eine Basis $\mathcal{A}$ des $\mathbb{R}^3$. Wir betrachten

$$x = \begin{pmatrix} -2 \\ 7 \\ 3 \end{pmatrix} = 31 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 11 \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} - 35 \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad x_{\mathcal{A}} = \begin{pmatrix} 31 \\ 11 \\ -35 \end{pmatrix}.$$

Dann gilt $x \neq x_{\mathcal{A}}$. □

Beispiel 7.3. Sei $m \in \mathbb{N}$ und $\mathcal{E} \subseteq \mathbb{R}^m$ die kanonische Basis des $\mathbb{R}^m$. Dann gilt $x = x_{\mathcal{E}}$ für alle $x \in \mathbb{R}^m$. □

Satz und Definition 7.4. *Seien $m, n \in \mathbb{N}$. Seien $\mathcal{A} = \{a_1, \ldots, a_m\}$ und $\mathcal{B} = \{b_1, \ldots, b_n\}$ Basen von X respektive Y. Sei $\varphi : X \to Y$ eine lineare Abbildung.*

(1) *Für alle $\mu = 1, \ldots, m$ gibt es eindeutig bestimmte Koeffizienten*

$$\gamma_{1\mu}, \ldots, \gamma_{n\mu} \in \mathbb{R}$$

derart, dass

$$\varphi(a_\mu) = \sum_{\nu=1}^{n} \gamma_{\nu\mu} b_\nu = \gamma_{1\mu} b_1 + \ldots + \gamma_{n\mu} b_n$$

gilt. Dabei wird zur Berechnung von $\varphi(a_\mu)$ über den ersten Index ν von $\gamma_{\nu\mu}$ summiert. Die $(n \times m)$-Matrix

$$M_{\mathcal{A}\mathcal{B}}(\varphi) = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nm} \end{pmatrix} \in \mathbb{R}^{n \times m}$$

heißt die Matrixdarstellung *von φ bezüglich $\mathcal{A}$ und $\mathcal{B}$.*

(2) *Für alle Vektoren $x \in X$ gilt*

$$(\varphi(x))_{\mathcal{B}} = M_{\mathcal{A}\mathcal{B}}(\varphi)\, x_{\mathcal{A}}\,.$$

Dabei sind $x_{\mathcal{A}}$ und $(\varphi(x))_{\mathcal{B}}$ die Koordinatenvektoren von x bezüglich $\mathcal{A}$ respektive von $\varphi(x)$ bezüglich $\mathcal{B}$.

(3) *Die Matrixdarstellung* $M_{\mathcal{AB}}(\varphi)$ *ist die einzige Matrix* $M \in \mathbb{R}^{n\times m}$, *die*

$$(\varphi(x))_{\mathcal{B}} = M x_{\mathcal{A}}$$

für alle $x \in X$ *erfüllt.*

(4) *Sei* $(M_{\mathcal{AB}}(\varphi))_\mu$ *die* μ*-te Spalte der Matrix* $M_{\mathcal{AB}}(\varphi)$. *Es gilt*

$$(M_{\mathcal{AB}}(\varphi))_\mu = (\varphi(a_\mu))_{\mathcal{B}} = \begin{pmatrix} \gamma_{1\mu} \\ \vdots \\ \gamma_{n\mu} \end{pmatrix}$$

für alle $\mu = 1, \ldots, m$.

(5) *Die Matrixdarstellung* $M_{\mathcal{AB}}(\varphi)$ *entsteht aus der Treppennormalform*

$$T(b_1, \ldots, b_n, (\varphi(a_1))_{\mathcal{B}}, \ldots, (\varphi(a_m))_{\mathcal{B}}) \in \mathrm{T}(\mathbb{R}^{n'\times(n+m)})$$

durch Streichen der $(n' - n)$ *Nullzeilen und der ersten* n *Spalten. Dabei ist*

$$T(b_1, \ldots, b_n, (\varphi(a_1))_{\mathcal{B}}, \ldots, (\varphi(a_m))_{\mathcal{B}})$$

die Treppennormalform der Matrix mit den $(n + m)$ *Spalten*

$$b_1, \ldots, b_n, (\varphi(a_1))_{\mathcal{B}}, \ldots, (\varphi(a_m))_{\mathcal{B}} \in \mathbb{R}^{n'} .$$

(6) *Im Spezialfall* $X = Y$ *mit* $\mathcal{A} = \mathcal{B}$ *und* $n = m$ *sowie* $a_\nu = b_\nu$ *für* $\nu = 1, \ldots, n$ *heißt* $M_{\mathcal{AA}}(\varphi)$ *die Matrixdarstellung von* φ *bezüglich* $\mathcal{A}$.

(7) *Im Spezialfall* $X = Y$ *mit* $n = m$ *und* $\varphi = \mathrm{id}_X$ *heißt die Matrixdarstellung*

$$M_{\mathcal{AB}} = M_{\mathcal{AB}}(\mathrm{id}_X)$$

die Transformationsmatrix *oder* Übergangsmatrix *von* $\mathcal{A}$ *nach* $\mathcal{B}$.

Wir heben die Stellung der Indizes in (1) *ausdrücklich hervor. In* (1) *wird über den ersten Index* ν *summiert, damit in* (2) *der Koordinatenvektor* $y_{\mathcal{B}}$ *durch linksseitige Multiplikation mit der Matrix* $M_{\mathcal{AB}}(\varphi)$ *aus dem Koordinatenvektor* $x_{\mathcal{A}}$ *hervorgeht.*

Beweis. Zu jedem Vektor $x \in X$ und jedem Vektor $y \in Y$ gibt es eindeutig bestimmte Koeffizienten $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ und $\beta_1, \ldots, \beta_n \in \mathbb{R}$ mit

$$x = \sum_{\mu=1}^{m} \alpha_\mu a_\mu , \quad y = \sum_{\nu=1}^{n} \beta_\nu b_\nu .$$

Wir betrachten für $x \in X$ den Vektor $\varphi(x) \in Y$. Es gilt

$$\varphi(x) = \sum_{\mu=1}^{m} \alpha_\mu \varphi(a_\mu) = \sum_{\mu=1}^{m} \left(\sum_{\nu=1}^{n} \alpha_\mu \gamma_{\nu\mu} \right) b_\nu = \sum_{\nu=1}^{n} \left(\sum_{\mu=1}^{m} \gamma_{\nu\mu} \alpha_\mu \right) b_\nu .$$

Folglich gilt $y = \varphi(x)$ genau dann, wenn

$$\beta_\nu = \sum_{\mu=1}^{m} \gamma_{\nu\mu}\alpha_\mu \,.$$

für alle $\nu = 1, \ldots, n$) gilt. Nach dieser Überlegung ist es klar, dass (2) bis (5) lediglich Umformulierungen von (1) sind. □

Beispiele 7.5.

(1) Sei $A \in \mathbb{R}^{n\times m}$ beliebig gegeben. Nach Satz 5.9 ist $\varphi_A : \mathbb{R}^m \to \mathbb{R}^n$ mit

$$\varphi(x) = Ax$$

eine lineare Abbildung. Nach Satz 7.4 gilt

$$A = M_{\mathcal{E}_m\mathcal{E}_n}(\varphi_A)\,.$$

(2) Für für alle Basen $\mathcal{A}$ von X gilt

$$M_{\mathcal{A}\mathcal{A}} = M_{\mathcal{A}\mathcal{A}}(\mathrm{id}_X) = E_m\,.$$

(3) Seien $\mathcal{A} = \{a_1, \ldots, a_m\}$ und $\mathcal{B} = \{b_1, \ldots, b_m\}$ zwei beliebige Basen des $\mathbb{R}^m$ sowie $A = (a_1, \ldots, a_m) \in \mathrm{GL}(m, \mathbb{R})$ und $B = (b_1, \ldots, b_m) \in \mathrm{GL}(m, \mathbb{R})$ die entsprechenden invertierbaren Matrizen mit den Spalten $a_1, \ldots, a_m$ respektive $b_1, \ldots, b_m$. Nach Satz 7.4 gilt

$$M_{\mathcal{A}\mathcal{B}} = M_{\mathcal{A}\mathcal{B}}(\mathrm{id}_{\mathbb{R}^m}) = B^{-1}A\,.$$

(4) Wenn $\varphi : X \to Y$ die Nullabbildung ist, dann gilt

$$M_{\mathcal{A}\mathcal{B}}(\varphi) = 0_{m\times n} \in \mathbb{R}^{n\times m}$$

für alle Basen $\mathcal{A}$ von X und alle Basen $\mathcal{B}$ von Y. □

Satz 7.6. *Seien $\mathcal{A}, \mathcal{B}, \mathcal{C}$ Basen von X, Y respektive Z. Seien $\varphi, \psi : X \to Y$ und $\eta : Y \to Z$ lineare Abbildungen. Sei $c \in \mathbb{R}$. Dann gelten die folgenden Regeln.*

(1) $M_{\mathcal{A}\mathcal{B}}(c\varphi) = cM_{\mathcal{A}\mathcal{B}}(\varphi)$.

(2) $M_{\mathcal{A}\mathcal{B}}(\varphi + \psi) = M_{\mathcal{A}\mathcal{B}}(\varphi) + M_{\mathcal{A}\mathcal{B}}(\psi)$.

(3) $M_{\mathcal{A}\mathcal{C}}(\eta \circ \varphi) = M_{\mathcal{B}\mathcal{C}}(\eta)M_{\mathcal{A}\mathcal{B}}(\varphi)$.

(4) $\dim(\mathrm{im}(\varphi)) = \dim(\mathrm{im}(M_{\mathcal{A}\mathcal{B}}(\varphi))$.

(5) $\dim(\ker(\varphi)) = \dim(\ker(M_{\mathcal{A}\mathcal{B}}(\varphi))$.

(6) *φ ist genau dann surjektiv, wenn* $\mathrm{rg}(M_{\mathcal{A}\mathcal{B}}(\varphi)) = \dim(Y)$ *gilt.*

(7) *φ ist genau dann injektiv, wenn* $\ker(M_{\mathcal{A}\mathcal{B}}(\varphi)) = \{0\}$ *gilt.*

Wir wenden uns nun linearen Abbildungen $\varphi : \mathbb{R}^m \to \mathbb{R}^m$ zu, die hinsichtlich einer geeigneten Basis $\mathcal{V}$ eine Diagonalmatrix als Matrixdarstellung besitzen.

Satz 7.7. *Sei* $\mathcal{E} = \{e_1, \dots, e_m\}$ *die kanonische Basis und* $\mathcal{V} = \{v_1, \dots, v_m\}$ *eine beliebige Basis des* $\mathbb{R}^m$. *Sei* $V = (v_1, \dots, v_m) \in \mathrm{GL}(m, \mathbb{R})$ *die invertierbare Matrix mit den Spalten* $v_1, \dots, v_m$. *Dann gelten für alle* $x \in \mathbb{R}^m$ *und alle* $\varphi \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^m)$ *die Formeln* (1), (2), (3).

(1) $x_{\mathcal{V}} = V^{-1} x$.

(2) $(\varphi(x))_{\mathcal{V}} = M_{\mathcal{V}\mathcal{V}}(\varphi)\, x_{\mathcal{V}}$.

(3) $M_{\mathcal{V}\mathcal{V}}(\varphi) = V^{-1} M_{\mathcal{E}\mathcal{E}}(\varphi)\, V$.

Beweis. Die Gültigkeit der Formeln (1) und (2) folgt direkt aus 7.1 und 7.4. Im Fall $\mathcal{V} = \mathcal{E}$ und $V = E$ gelten (1), (2), (3) trivialerweise. Seien $\varphi \in \mathrm{L}(\mathbb{R}^m, \mathbb{R}^m)$ und $x \in \mathbb{R}^m$ beliebig gewählt. Dann gilt

$$\begin{aligned} M_{\mathcal{V}\mathcal{V}}(\varphi)\, V^{-1} x &= M_{\mathcal{V}\mathcal{V}}(\varphi)\, x_{\mathcal{V}} \\ &= (\varphi(x))_{\mathcal{V}} \\ &= V^{-1}\, \varphi(x) \\ &= V^{-1}\, (\varphi(x))_{\mathcal{E}} \\ &= V^{-1} M_{\mathcal{E}\mathcal{E}}(\varphi)\, x_{\mathcal{E}} \\ &= V^{-1} M_{\mathcal{E}\mathcal{E}}(\varphi)\, x\,. \end{aligned}$$

Weil $x \in \mathbb{R}^m$ beliebig gewählt war, folgt

$$M_{\mathcal{V}\mathcal{V}}(\varphi)\, V^{-1} = V^{-1} M_{\mathcal{E}\mathcal{E}}(\varphi)\,.$$

Multiplikation von rechts mit V ergibt (2). Damit ist der Beweis beendet. $\square$

8 Diagonalisierung

Sei $m \in \mathbb{N}$. Gegeben sei eine lineare Abbildung $\varphi : \mathbb{R}^m \to \mathbb{R}^m$ mit der Matrixdarstellung

$$A = M_{\mathcal{E}\mathcal{E}}(\varphi) \in \mathbb{R}^{m\times m} .$$

Gesucht ist eine Basis $\mathcal{V} = \{v_1 \dots, v_m\}$ des $\mathbb{R}^m$ derart, dass die Matrixdarstellung

$$M_{\mathcal{V}\mathcal{V}}(\varphi) = V^{-1}AV$$

möglichst einfach wird. Dabei ist $V = (v_1 \dots, v_m) \in \mathbb{R}^{m\times m}$ die Matrix mit den Spalten $v_1 \dots, v_m$. Wir nehmen an, dass wir $\mathcal{V}$ so wählen können, dass $M_{\mathcal{V}\mathcal{V}}(\varphi)$ eine *Diagonalmatrix* ist.

Definition 8.1. *Eine Matrix $D = (d_{\nu\mu}) \in \mathbb{R}^{m\times m}$ heißt* Diagonalmatrix*, wenn $d_{\nu\mu} = 0$ für alle $\nu, \mu \in \{1, \dots, m\}$ mit $\nu \neq \mu$ gilt. Wir schreiben dann*

$$D = \operatorname{diag}(\lambda_1, \dots, \lambda_m) ,$$

wobei wir $\lambda_\mu = d_{\mu\mu}$ für $\mu = 1, \dots, m$ gesetzt haben.

Die zusätzliche Annahme, dass $M_{\mathcal{V}\mathcal{V}}(\varphi)$ eine Diagonalmatrix ist, führt auf den Begriff des *Eigenwertes* und des *Eigenvektors* einer Matrix.

Satz 8.2. *Sei $m \in \mathbb{N}$. Für $A \in \mathbb{R}^{m\times m}$ und $V = (v_1, \dots, v_m) \in \mathrm{GL}(m, \mathbb{R})$ sind die folgenden Aussagen* (1) *bis* (4) *äquivalent.*

(1) $\operatorname{diag}(\lambda_1, \dots, \lambda_m) = V^{-1}AV$.

(2) $V \circ \operatorname{diag}(\lambda_1, \dots, \lambda_m) = A \circ V$.

(3) $A \circ V = (\lambda_1 v_1, \dots, \lambda_m v_m)$.

(4) $(\forall \mu = 1, \dots, m) : \quad Av_\mu = \lambda_\mu v_\mu$.

Definition 8.3. *Seien $m \in \mathbb{N}$ und $A \in \mathbb{R}^{m\times m}$ gegeben.*

(1) *Ein Vektor $v \in \mathbb{R}^m \setminus \{0\}$ heißt ein* Eigenvektor *von A genau dann, wenn es ein $\lambda \in \mathbb{R}$ mit $Av = \lambda v$ gibt. Die reelle Zahl λ heißt der* Eigenwert *von A zum Eigenvektor v.*

(2) *Die Matrix A heißt* diagonalisierbar*, wenn es eine Basis $\mathcal{V} = \{v_1, \dots, v_m\}$ des $\mathbb{R}^m$ aus Eigenvektoren $v_1, \dots, v_m$ von A gibt.*

(3) *Wenn die Matrix A diagonalisierbar ist, dann gibt es linear unabhängige Eigenvektoren $v_1, \dots, v_m \in \mathbb{R}^m$. Seien $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ die zugehörigen Eigenwerte. Sei $V = (v_1, \dots, v_m) \in \mathrm{GL}(m, \mathbb{R})$. Die Beziehung*

$$\operatorname{diag}(\lambda_1, \dots, \lambda_m) = V^{-1}AV$$

heißt eine Diagonaldarstellung *von A.*

Beispiel 8.4. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $v_1, v_2, v_3 \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Wir zeigen, dass A diagonalisierbar ist. Die Menge $\mathcal{V} = \{v_1, v_2, v_3\}$ ist eine Basis des $\mathbb{R}^3$ aus Eigenvektoren von A. Aus

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix} \begin{pmatrix} 1 & -3 & 0 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -6 & 0 \\ 2 & 2 & 1 \\ 4 & 2 & 2 \end{pmatrix}$$

folgt, dass $\lambda_1 = \lambda_2 = 2$ und $\lambda_3 = 1$. Siehe Satz 8.2. Zum Nachweis, dass $\mathcal{V}$ eine Basis ist, zeigen wir, dass die Matrix $V = (v_1, v_2, v_3)$ invertierbar ist. Siehe Satz 6.14.

$$\begin{array}{rrr|rrr|l}
1 & -3 & 0 & 1 & 0 & 0 & \\
1 & 1 & 1 & 0 & 1 & 0 & \\
2 & 1 & 2 & 0 & 0 & 1 & \\
\hline
1 & -3 & 0 & 1 & 0 & 0 & \\
0 & 4 & 1 & -1 & 1 & 0 & Z_2 - Z_1 \\
0 & 7 & 2 & -2 & 0 & 1 & Z_3 - 2Z_1 \\
\hline
1 & -3 & 0 & 1 & 0 & 0 & \\
0 & 4 & 1 & -1 & 1 & 0 & \\
0 & 3 & 1 & -1 & -1 & 1 & Z_3 - Z_2 \\
\hline
1 & 0 & 1 & 0 & -1 & 1 & Z_1 + Z_3 \\
0 & 1 & 0 & 0 & 2 & -1 & Z_2 - Z_3 \\
0 & 3 & 1 & -1 & -1 & 1 & \\
\hline
1 & 0 & 1 & 0 & -1 & 1 & \\
0 & 1 & 0 & 0 & 2 & -1 & \\
0 & 0 & 1 & -1 & -7 & 4 & Z_3 - 3Z_2 \\
\hline
1 & 0 & 0 & 1 & 6 & -3 & Z_1 - Z_3 \\
0 & 1 & 0 & 0 & 2 & -1 & \\
0 & 0 & 1 & -1 & -7 & 4 &
\end{array}$$

Wir erhalten die Diagonaldarstellung

$$V^{-1}AV = \begin{pmatrix} 1 & 6 & -3 \\ 0 & 2 & -1 \\ -1 & -7 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix} \begin{pmatrix} 1 & -3 & 0 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix} = \operatorname{diag}(2, 2, 1).$$

Siehe Beispiel 8.6. $\square$

Satz 8.5. *Seien $m \in \mathbb{N}$ und eine diagonalisierbare Matrix $A \in \mathbb{R}^{m \times m}$ gegeben. Seien $\mathcal{V} = \{v_1, \ldots, v_m\}$ eine Basis des $\mathbb{R}^m$ aus Eigenvektoren $v_1, \ldots, v_m$ der Matrix A und $V = (v_1, \ldots, v_m) \in \mathrm{GL}(m, \mathbb{R})$ die invertierbare Matrix mit den Spalten $v_1, \ldots, v_m$. Weiter seien $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ die Eigenwerte von A mit $Av_\mu = \lambda_\mu v_\mu$ für $\mu = 1, \ldots, m$.*

(1) *Es gilt die Diagonaldarstelllung*

$$A = V \circ \mathrm{diag}(\lambda_1, \ldots, \lambda_m) \circ V^{-1} .$$

(2) *Sei $n \in \mathbb{N}_0$. Für ein reelles Polynom*

$$p(\lambda) = \sum_{\nu=0}^{n} \alpha_\nu \lambda^\nu = \alpha_n \lambda^n + \ldots + \alpha_1 \lambda + \alpha_0 \in \mathbb{R}[\lambda]$$

mit $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$ sei $p(A) \in \mathbb{R}^{m \times m}$ die Matrix mit

$$p(A) = \sum_{\nu=0}^{n} \alpha_\nu A^\nu = \alpha_n A^n + \ldots + \alpha_1 A + \alpha_0 E_m .$$

Dabei definieren wir die Potenzen der Matrix A induktiv durch $A^0 = E_m$ und $A^{k+1} = A \circ A^k$ für alle $k \in \mathbb{N}_0$.

(3) *Für jedes reelle Polynom $p(\lambda) \in \mathbb{R}[\lambda]$ gilt*

$$p(A) = V \circ \mathrm{diag}(p(\lambda_1), \ldots, p(\lambda_m)) \circ V^{-1} .$$

Für $p(\lambda) = \lambda$ erhalten wir die Diagonaldarstellung (1) zurück.

Beispiel 8.6. Nach Beispiel 8.4 ist die Matrix $A \in \mathbb{R}^{3 \times 3}$ mit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix}$$

diagonalisierbar. Es gilt

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix} = \begin{pmatrix} 1 & -3 & 0 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 6 & -3 \\ 0 & 2 & -1 \\ -1 & -7 & 4 \end{pmatrix} .$$

Mit dieser Diagonaldarstellung

$$A = V \circ \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) \circ V^{-1}$$

können Polynome $p(A)$ berechnet werden. Für alle $p(\lambda) \in \mathbb{R}[\lambda]$ gilt

$$p(A) = V \circ \mathrm{diag}(p(\lambda_1), \ldots, p(\lambda_m)) \circ V^{-1} .$$

Siehe Satz 8.5.

(1) Zuerst betrachten wir das Polynom

$$\begin{aligned}\chi(\lambda) &= (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)\\ &= (\lambda-2)^2(\lambda-1) = \lambda^3 - 5\lambda^2 + 8\lambda - 4\,.\end{aligned}$$

Es gilt

$$\begin{aligned}\chi(A) &= V \circ \operatorname{diag}(\chi(2), \chi(2), \chi(1)) \circ V^{-1}\\ &= V \circ \operatorname{diag}(0,0,0) \circ V^{-1} = \operatorname{diag}(0,0,0)\\ &= \begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\end{aligned}$$

Siehe Beispiel 14.2 und Satz 14.4.

(2) Nun betrachten wir das Polynom

$$q(\lambda) = \lambda^3 - 5\lambda^2 - 4 = \chi(\lambda) - 8\lambda\,.$$

Nach Teil (1) gilt

$$q(A) = \chi(A) - 8A = -8A\,. \tag{8.1}$$

Wir berechnen $q(A)$ mit Hilfe der Diagonaldarstellung von A. Die Identität (8.1) kann zur Probe verwendet werden. Es gilt

$$\begin{aligned}A^3 - 5A^2 - 4E_3 &= \begin{pmatrix} 1 & -3 & 0\\ 1 & 1 & 1\\ 2 & 1 & 2\end{pmatrix}\begin{pmatrix} q(2) & 0 & 0\\ 0 & q(2) & 0\\ 0 & 0 & q(1)\end{pmatrix}\begin{pmatrix} 1 & 6 & -3\\ 0 & 2 & -1\\ -1 & -7 & 4\end{pmatrix}\\ &= \begin{pmatrix} 1 & -3 & 0\\ 1 & 1 & 1\\ 2 & 1 & 2\end{pmatrix}\begin{pmatrix} -16 & 0 & 0\\ 0 & -16 & 0\\ 0 & 0 & -8\end{pmatrix}\begin{pmatrix} 1 & 6 & -3\\ 0 & 2 & -1\\ -1 & -7 & 4\end{pmatrix}\\ &= \begin{pmatrix} 1 & -3 & 0\\ 1 & 1 & 1\\ 2 & 1 & 2\end{pmatrix}\begin{pmatrix} -16 & -96 & 48\\ 0 & -32 & 16\\ 8 & 56 & -32\end{pmatrix}\\ &= \begin{pmatrix} -16 & 0 & 0\\ -8 & -72 & 32\\ -16 & -112 & 48\end{pmatrix} = (-8)\begin{pmatrix} 2 & 0 & 0\\ 1 & 9 & -4\\ 1 & 14 & -6\end{pmatrix} = -8A\,.\end{aligned}$$

Natürlich können die Matrizen $\chi(A)$ und $q(A)$ auch direkt nach der Definition (2) in 8.5 berechnet werden. □

Es gibt allerdings quadratische Matrizen, die nicht diagonalisierbar sind.

Beispiel 8.7. Die reelle (2×2)-Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

ist nicht diagonalisierbar. Aus

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

folgt $y = 0$. Daher kann es keine Basis des $\mathbb{R}^2$ aus Eigenvektoren der Matrix A geben. $\square$

In Abschnitt 12 werden wir beweisen, dass eine quadratische Marix genau dann invertierbar ist, wenn ihre Determinante von Null verschieden ist. Siehe Satz 12.5.

Satz und Definition 8.8. *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$.*

(1) *Das reelle Polynom $\chi_A(\lambda) = \det(A - \lambda E_n)$ heißt das* charakteristische Polynom *von A.*

(2) *$\lambda \in \mathbb{R}$ ist genau dann ein Eigenwert von A, wenn $\chi_A(\lambda) = 0$ gilt.*

Beispiel 8.9. Nach 8.4 hat die Matrix

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$$

den einfachen Eigenwert 1 und den doppelten Eigenwert 2. Wir rechnen dies nach, indem wir das charakteristische Polynom $\chi_A(\lambda)$ von A nach der rekursiven Definition 1.4 auswerten.

$$\begin{aligned} \chi_A(\lambda) &= (2 - \lambda) \cdot \big((9 - \lambda)(-6 - \lambda) + 56\big) - 1 \cdot 0 + 2 \cdot 0 \\ &= (2 - \lambda)(\lambda^2 - 3\lambda - 2) \\ &= (-1)(\lambda - 1)(\lambda - 2)^2 . \end{aligned}$$

Wir erhalten $\det(A) = \chi_A(0) = 4$. Siehe Beispiel 1.5. $\square$

9 Bestandteile der strikten Bruhat-Zerlegung

Nach dem Zerlegungssatz 6.15 lässt sich jede invertierbare Matrix als Produkt von endlich vielen Elementarmatizen schreiben. Solche eine Zerlegung ist nicht eindeutig bestimmt. Es gibt Zerlegungen in Elementarmatrizen, deren Faktoren sich in vier Matrizen zusammenfassen lassen, wobei jede dieser vier Matrizen eine bestimmte Bauart besitzt. Nach Satz 10.1 besitzt jede Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ eine Zerlegung der Form

$$A = L \circ D \circ P \circ U\,,$$

wobei L eine untere Dreiecksmatrix, D eine Diagonalmatrix, P eine Permutationsmatrix und U eine obere Dreiecksmatrix ist. Dabei sind die Diagonalelemente von L und U auf 1 normiert. Solch eine Zerlegung heißt eine *Bruhat-Zerlegung*. Die Matrizen D und P sind in jedem Fall eindeutig bestimmt. Die normierten Dreiecksmatrizen L und U sind eindeutig bestimmt, wenn

$$P \circ U \circ P^{-1}$$

ebenfalls eine normierte obere Dreiecksmatrix ist. Solche eine eindeutig bestimmte Zerlegung heißt eine *strikten Bruhat-Zerlegung von A*. In diesem Abschnitt beschreiben wird die Bestandteile dieser Zerlegung.

Beispiel 9.1. Wir betrachten noch einmal die invertierbare reelle (2×2)-Matrix aus Beispiel 6.16.

$$\begin{aligned}
\begin{pmatrix} \frac{1}{5} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{7} \end{pmatrix} &= \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{6} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{252} \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ \frac{5}{6} & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{252} \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ \frac{5}{6} & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{252} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{6} \\ 0 & 1 \end{pmatrix}\,.
\end{aligned}$$

In diesem Fall ist die Permutationsmatrix P die Einheitsmatrix E_2. Daher ist die Zerlegung in der letzten Zeile die strikte Bruhat-Zerlegung. □

In 9.2 und 9.8 beschreiben wir die Bestandteile der strikten Bruhat-Zerlegung. Die Erörterungen dieses Abschnittes bereiten den Beweis des Satzes 10.1 vor.

Satz und Definition 9.2. *Sei* $n \in \mathbb{N}$.

(1) *Sei* $\mathrm{D}(n, \mathbb{R}) = \{D = (d_{\nu\mu}) \in \mathbb{R}^{n\times n} \mid d_{\nu\mu} = 0,\ \nu \neq \mu,\ \nu, \mu = 1, \ldots, n\}$ *die Menge der reellen* $(n \times n)$*-Diagonalmatrizen. Dann gelten:*

(1.1) $\mathrm{D}^{\times}(n, \mathbb{R}) = \{\mathrm{diag}(d_1, \ldots, d_n) \mid d_1, \ldots d_n \in \mathbb{R}^{\times}\} \subseteq \mathrm{GL}(n, \mathbb{R})$.

(1.2) $E_n = \mathrm{diag}(1, \ldots, 1) \in \mathrm{GL}(n, \mathbb{R})$.

(2) *Sei* $\mathrm{L}^{\llcorner}(n,\mathbb{R}) = \{(l_{\nu\mu}) \in \mathbb{R}^{n\times n} \mid l_{\nu\mu} = 0,\, \mu > \nu\}$ *die Menge der unteren Dreiecksmatrizen. Dann gelten:*

(2.1) $\mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R}) = \{(l_{\nu\mu}) \in \mathrm{L}^{\llcorner}(n,\mathbb{R}) \mid l_{\nu\nu} \in \mathbb{R}^{\times},\, \nu = 1,\dots,n\} \subseteq \mathrm{GL}(n,\mathbb{R})$.

(2.2) $\mathrm{L}^{\llcorner}_{1}(n,\mathbb{R}) = \{(l_{\nu\mu}) \in \mathrm{L}^{\llcorner}(n,\mathbb{R}) \mid l_{\nu\nu} = 1,\, \nu = 1,\dots,n\} \subseteq \mathrm{GL}(n,\mathbb{R})$.

(3) *Sei* $\mathrm{U}^{\urcorner}(n,\mathbb{R}) = \{(u_{\nu\mu}) \in \mathbb{R}^{n\times n} \mid u_{\nu\mu} = 0,\, \nu > \mu\}$ *die Menge der oberen Dreiecksmatrizen. Dann gelten:*

(3.1) $\mathrm{U}^{\urcorner}_{\times}(n,\mathbb{R}) = \{(u_{\nu\mu}) \in \mathrm{U}^{\urcorner}(n,\mathbb{R}) \mid u_{\nu\nu} \in \mathbb{R}^{\times},\, \nu = 1,\dots,n\} \subseteq \mathrm{GL}(n,\mathbb{R})$.

(3.2) $\mathrm{U}^{\urcorner}_{1}(n,\mathbb{R}) = \{(u_{\nu\mu}) \in \mathrm{U}^{\urcorner}(n,\mathbb{R}) \mid u_{\nu\nu} = 1,\, \nu = 1,\dots,n\} \subseteq \mathrm{GL}(n,\mathbb{R})$.

(4) *Sei* $X \subseteq \mathrm{GL}(n,\mathbb{R})$ *eine der fünf Teilmengen* $\mathrm{D}^{\times}(n,\mathbb{R})$, $\mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R})$, $\mathrm{L}^{\llcorner}_{1}(n,\mathbb{R})$, $\mathrm{U}^{\urcorner}_{\times}(n,\mathbb{R})$, $\mathrm{U}^{\urcorner}_{1}(n,\mathbb{R})$. *Dann gelten:*

(4.1) $E_n \in X$.

(4.2) $(\forall A, B \in X) : A \circ B \in X$.

(4.3) *Die Diagonalelemente verhalten sich bei Produktbildung in* X *multiplikativ. Seien* $A = (a_{\nu\mu})$, $B = (b_{\mu\lambda})$, $C = (c_{\nu\lambda}) \in X$ *Matrizen mit* $C = AB$. *Dann gilt*

$$c_{\nu\nu} = a_{\nu\nu} b_{\nu\nu}$$

für alle $\nu = 1,\dots,n$.

(4.4) $(\forall A \in X) : A^{-1} \in X$.

(4.5) $\mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R}) \cap \mathrm{U}^{\urcorner}_{\times}(n,\mathbb{R}) = \mathrm{D}^{\times}(n,\mathbb{R})$.

(4.6) $\mathrm{L}^{\llcorner}_{1}(n,\mathbb{R}) \cap \mathrm{U}^{\urcorner}_{1}(n,\mathbb{R}) = \{E_n\}$.

Siehe Satz und Definition 9.8*, Aussage* (11)*.*

Wir heben ausdrücklich hervor, dass die Produktregel (4.3) *für beliebige Matrizen aus* $\mathbb{R}^{n\times n}$ *im Allgemeinen nicht gilt.*

Beweis. Wir beweisen (4.2), (4.3) und (4.4) im Fall $X = \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R})$. Alle anderen Aussagen des Satzes sind klar oder ergeben sich aus diesem Spezialfall. Für alle $M \in \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R})$ gilt $T(M) = E_n$. Also gilt $\mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R}) \subseteq \mathrm{GL}(n,\mathbb{R})$. Das Verfahren von Gauß-Jordan liefert außerdem $M^{-1} \in \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R})$. Siehe Satz 6.14. Seien

$$A = (a_{\nu\mu}) \in \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R}), \quad B = (b_{\mu\lambda}) \in \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R}), \quad C = (c_{\nu\lambda}) \in \mathbb{R}^{n\times n}$$

mit $C = AB$ gegeben. Dann gilt $C \in \mathrm{GL}(n,\mathbb{R})$. Weiter gilt $a_{\nu\mu} \neq 0$ allenfalls für $\mu \leq \nu$. Entsprechend gilt $b_{\mu\lambda} \neq 0$ allenfalls für $\lambda \leq \mu$. Folglich gilt

$$c_{\nu\lambda} = \sum_{\mu=1}^{n} a_{\nu\mu} b_{\mu\lambda} \neq 0$$

allenfalls für $\lambda \leq \nu$. Dabei gilt

$$c_{\nu\lambda} = \sum_{\lambda\leq\mu\leq\nu} a_{\nu\mu} b_{\mu\lambda}\,.$$

Nun folgen

$$C = AB \in \mathrm{L}^{\llcorner}(n,\mathbb{R}) \cap \mathrm{GL}(n,\mathbb{R}) = \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R})\,, \quad c_{\nu\nu} = a_{\nu\nu} b_{\nu\nu}$$

für alle $\nu = 1, \ldots, n$. Damit ist der Beweis beendet. $\square$

Beispiel 9.3. Das Produkt zweier unterer Dreiecksmatrizen aus $\mathrm{L}^{\llcorner}_{1}(3,\mathbb{R})$ ist nach 9.2 Aussage (4.2) eine untere Dreiecksmatrix aus $\mathrm{L}^{\llcorner}_{1}(3,\mathbb{R})$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 6 & 7 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 0 \\ 29 & 11 & 1 \end{pmatrix}.$$

Nach 9.2 Aussage (2.2) ist eine untere Dreiecksmatrix aus $\mathrm{L}^{\llcorner}_{1}(3,\mathbb{R})$ invertierbar. Die inverse Matrix ist nach 9.2 Aussage (4.4) ebenfalls eine untere Dreiecksmatrix aus $\mathrm{L}^{\llcorner}_{1}(3,\mathbb{R})$. Das Verfahren von Gauß-Jordan liefert

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{pmatrix}.$$

Die Einzelheiten sind in folgendem Schema enthalten.

$$\begin{array}{ccc|rrr|l}
1 & 0 & 0 & 1 & 0 & 0 & \\
2 & 1 & 0 & 0 & 1 & 0 & \\
3 & 4 & 1 & 0 & 0 & 1 & \\
\hline
1 & 0 & 0 & 1 & 0 & 0 & Z_1 \\
0 & 1 & 0 & -2 & 1 & 0 & Z_2 - 2Z_1 \\
0 & 4 & 1 & -3 & 0 & 1 & Z_3 - 3Z_1 \\
\hline
1 & 0 & 0 & 1 & 0 & 0 & Z_1 \\
0 & 1 & 0 & -2 & 1 & 0 & Z_2 \\
0 & 0 & 1 & 5 & -4 & 1 & Z_3 - 4Z_2
\end{array}$$

Siehe Satz 6.14. Wir machen die Probe. Offenbar gilt

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nach 9.2 Aussage (4.3) verhalten sich die Diagonalelemente bei Produktbildung in $\mathrm{L}^{\llcorner}_{\times}(3,\mathbb{R})$ multiplikativ.

$$\begin{pmatrix} 9 & 0 & 0 \\ -3 & 4 & 0 \\ 5 & 1 & -7 \end{pmatrix} \begin{pmatrix} -3 & 0 & 0 \\ 2 & 5 & 0 \\ 4 & 3 & 5 \end{pmatrix} = \begin{pmatrix} -27 & 0 & 0 \\ 17 & 20 & 0 \\ -41 & -16 & -35 \end{pmatrix}.$$

$\square$

Der Beweis des Satz 10.1 ist konstruktiv. Die Konstruktion der normierten unteren Dreiecksmatrix $L \in L_{\mathbb{1}}^{\llcorner}(n,\mathbb{R})$ erfolgt schrittweise durch Konstruktion der $(n-1)$ ersten Spalten. Dabei ist die n-te Spalte gleich dem n-ten kanonischen Basisvektor e_n. Die einzelnen Schritte lassen sich durch Elementarmatrizen beschrieben. Das Produkt dieser Elementarmatrizen ergibt schließlich die gewünschte Matrix L. Zum Nachweis benötigen wir den folgenden Satz. Beispiele werden in 9.5 und 10.4 gegeben.

Satz 9.4. *Sei $n \in \mathbb{N}$. Für alle $a \in \mathbb{R}^n$ und $L \in \mathrm{L}_{\mathbb{1}}^{\llcorner}(n,\mathbb{R})$ gilt*

$$\begin{pmatrix} 1 & 0_{1\times n} \\ a & L \end{pmatrix} = \begin{pmatrix} 1 & 0_{1\times n} \\ a & E_n \end{pmatrix} \begin{pmatrix} 1 & 0_{1\times n} \\ 0_m & L \end{pmatrix} \in \mathrm{L}_{\mathbb{1}}^{\llcorner}(n+1,\mathbb{R})\,.$$

Die beiden Matrizen auf der rechten Seiten sind im Allgemeinen nicht vertauschbar. Die behauptete Identität lässt sich leicht durch Ausmultipizieren nachprüfen.

Beispiel 9.5.

(1) In $\mathrm{L}_{\mathbb{1}}^{\llcorner}(3,\mathbb{R})$ gilt

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix}.$$

Die beiden Matrizen auf der rechten Seite sind nicht vertauschbar.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 11 & 4 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix}.$$

(2) In $\mathrm{L}_{\mathbb{1}}^{\llcorner}(3,\mathbb{R})$ gilt

$$\begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{32} & 1 \end{pmatrix}.$$

Die beiden Matrizen auf der rechten Seite sind im Allgemeinen nicht vertauschbar. Siehe (1).

(3) In $\mathrm{L}_1^{\llcorner}(4,\mathbb{R})$ gilt

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{42} & l_{43} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{42} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & l_{43} & 1 \end{pmatrix}.$$

Dabei haben wir Satz 9.4 zweimal angewendet. Die drei Matrizen auf der rechten Seite sind im Allgemeinen nicht vertauschbar. Es ist also auf die Reihenfolge bei der Produktbildung zu achten. Siehe Beweis von Satz 10.1.

$\square$

Satz 9.6. *Seien m, $n \in \mathbb{N}$.*

(1) *Für alle $a \in \mathbb{R}^n$ gilt*

$$\begin{pmatrix} 1 & 0_{1\times n} \\ a & E_n \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0_{1\times n} \\ -a & E_n \end{pmatrix} \in \mathrm{L}_1^{\llcorner}(n+1,\mathbb{R}).$$

(2) *Für alle $a \in \mathbb{R}^n$ gilt*

$$\begin{pmatrix} E_m & 0 & 0_{m\times n} \\ 0_{1\times m} & 1 & 0_{1\times n} \\ 0_{n\times m} & a & E_n \end{pmatrix}^{-1} = \begin{pmatrix} E_m & 0 & 0_{m\times n} \\ 0_{1\times m} & 1 & 0_{1\times n} \\ 0_{n\times m} & -a & E_n \end{pmatrix} \in \mathrm{L}_1^{\llcorner}(m+n+1,\mathbb{R}).$$

Die behaupteten Identitäten folgen unmittelbar aus dem Verfahren von Gauß-Jordan. Sie lassen sich leicht durch Ausmultiplizieren nachprüfen.

Beispiel 9.7. Nach Beispiel 9.3 gilt

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{pmatrix}.$$

Wir können dies auch mit Hilfe von 9.4 und 9.6 ausrechnen. Aus der Zerlegung

$$\begin{pmatrix}1&0&0\\2&1&0\\3&4&1\end{pmatrix} = \begin{pmatrix}1&0&0\\2&1&0\\3&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&4&1\end{pmatrix}$$

erhalten wir

$$\begin{aligned}\begin{pmatrix}1&0&0\\2&1&0\\3&4&1\end{pmatrix}^{-1} &= \begin{pmatrix}1&0&0\\0&1&0\\0&4&1\end{pmatrix}^{-1}\begin{pmatrix}1&0&0\\2&1&0\\3&0&1\end{pmatrix}^{-1}\\ &= \begin{pmatrix}1&0&0\\0&1&0\\0&-4&1\end{pmatrix}\begin{pmatrix}1&0&0\\-2&1&0\\-3&0&1\end{pmatrix}\\ &= \begin{pmatrix}1&0&0\\-2&1&0\\5&-4&1\end{pmatrix}.\end{aligned}$$

Die vorletzte Zeile ist die matrizentheoretische Version der Kommentarspalte des Rechenschemas in Beispiel 9.3. □

Satz und Definition 9.8. *Seien $n, m \in \mathbb{N}$.*

(1) *Eine Abbildung $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ heißt eine* Permutation *oder eine* Vertauschung *der Zahlen $1, \ldots, n$, wenn es zu jedem $j \in \{1, \ldots, n\}$ genau ein $i \in \{1, \ldots, n\}$ mit $\pi(i) = j$ gibt. Sei S_n die Menge der Permutationen der Zahlen $1, \ldots, n$. Wir notieren eine Permutation $\pi \in \mathrm{S}_n$ als $(2 \times n)$-Matrix in der Form*

$$\pi = \begin{pmatrix}1 & \ldots & n\\ \pi(1) & \ldots & \pi(n)\end{pmatrix}.$$

Es gelten:

(1.1) $\mathrm{id}_{\{1,\ldots,n\}} \in \mathrm{S}_n$.

(1.2) $(\forall \pi, \sigma \in \mathrm{S}_n) : \pi \circ \sigma \in \mathrm{S}_n$.

(1.3) *Jede Permutation $\pi \in \mathrm{S}_n$ besitzt eine Umkehrfunktion $\pi^{-1} \in \mathrm{S}_n$.*

(1.4) $\#(\mathrm{S}_n) = n! = 1 \cdot \ldots \cdot n$.

(2) *Eine Permutation $\pi \in \mathrm{S}_n$ mit $\pi(i) = j$ und $\pi(j) = i$ und $\pi(\nu) = \nu$ für $\nu \in \{1, \ldots, n\}$ mit $\nu \neq i, j$ bezeichnen wir mit $\pi_{i \leftrightarrow j}$. Im Fall $n \geq 2$ und $i \neq j$ heißt $\pi_{i \leftrightarrow j}$ eine* Transposition.

(3) *Sei $n \geq 2$. Jede Permutation aus S_n ist eine Komposition endlich vieler Transpositionen aus S_n. Die identische Abbildung $\mathrm{id}_{\{1,\ldots,n\}}$ lässt sich als Produkt zweier Transpositionen schreiben.*

(4) *Eine Matrix* $P \in \mathbb{R}^{n\times n}$ *heißt eine* Permutationsmatrix, *wenn jede Zeile und jede Spalte von* P *genau einen Eintrag* 1 *enthält und alle anderen Einträge von* P *verschwinden, das heißt, gleich* 0 *sind. Sei* $\mathrm{Perm}(n,\mathbb{R})$ *die Menge der Permutationsmatrizen* $P \in \mathbb{R}^{n\times n}$. *Es gelten:*

(4.1) $E_n = (e_1, \ldots, e_n) \in \mathrm{Perm}(n,\mathbb{R})$.

(4.2) $(\forall P, Q \in \mathrm{Perm}(n,\mathbb{R}) : P \circ Q \in \mathrm{Perm}(n,\mathbb{R})$.

(4.3) $\mathrm{Perm}(n,\mathbb{R}) \subseteq \mathrm{GL}(n,\mathbb{R})$.

(4.4) $(\forall P \in \mathrm{Perm}(n,\mathbb{R}) : P^{-1} = P^t \in \mathrm{Perm}(n,\mathbb{R})$.

(4.5) $\#(\mathrm{Perm}(n,\mathbb{R})) = n!$.

(5) *Einer Permutation* $\pi \in \mathrm{S}_n$ *ordnen wir die Permutationsmatrix*

$$P(\pi) = (e_{\pi(1)}, \ldots, e_{\pi(n)}) \in \mathrm{Perm}(n,\mathbb{R})$$

zu. Also ist die μ*-te Spalte von* $P(\pi)$ *gleich dem* $\pi(\mu)$*-ten kanonischen Basisvektor* $e_{\pi(\mu)} \in \mathbb{R}^n$. *Es gilt*

$$P(\pi)e_\mu = e_{\pi(\mu)}$$

für alle $\mu = 1, \ldots, m$.

(6) *Zu jeder Permutationsmatrix* $P \in \mathrm{Perm}(n,\mathbb{R})$ *gibt es genau eine Permutation* $\pi \in \mathrm{S}_n$ *mit* $P = P(\pi)$. *Es gelten:*

(6.1) $P(\mathrm{id}_{1,\ldots,n}) = E_n$.

(6.2) $(\forall \pi, \sigma \in \mathrm{S}_n) : P(\pi \circ \sigma) = P(\pi) \circ P(\sigma)$.

(6.3) $(\forall \pi \in \mathrm{S}_n) : (P(\pi))^{-1} = P(\pi^{-1}) = (P(\pi))^t$.

(6.4) *Es gilt* $P(\pi_{i\leftrightarrow j}) = E_{i\leftrightarrow j}$ *für alle* $i, j \in \{1, \ldots, n\}$.

(7) *Sei* $\pi \in \mathrm{S}_n$. *Dann gilt*

$$P(\pi) = (p_{\nu\mu}), \quad p_{\nu\mu} = \delta_{\nu\pi(\mu)} = \delta_{\pi^{-1}(\nu)\mu}$$

für alle $\nu, \mu = 1, \ldots, n$. *Folglich ist die* $\pi(\nu)$*-te Zeile von* $P(\pi)$ *gleich der* ν*-ten Zeile der Einheitsmatrix* E_n. *Also gilt*

$$P(\pi) = \begin{pmatrix} e^t_{\pi^{-1}(1)} \\ \vdots \\ e^t_{\pi^{-1}(n)} \end{pmatrix}.$$

Siehe (6.3). *Die Formeln in* (5) *und* (7) *beschreiben die Spalten, Einträge respektive Zeilen der Permutationsmatrizen* $P(\pi)$. *Die Formeln in* (8) *und* (9) *beschreiben die Wirkung von* $P(\pi)$ *bei einer multiplikativen Anwendung von links respektive rechts.*

(8) *Seien* $\pi \in \mathrm{S}_n$ *und* $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times m}$. *Dann gilt*

$$P(\pi) \circ A = (\alpha_{\nu\mu}), \quad \alpha_{\nu\mu} = a_{\pi^{-1}(\nu)\mu}$$

für alle $\nu = 1, \ldots, n$ und alle $\mu = 1, \ldots, m$. Folglich ist die $\pi(\nu)$-te Zeile von $P(\pi) \circ A$ gleich der ν-ten Zeile von A. Sind $z_1^t, \ldots, z_n^t \in \mathbb{R}^m$ die Spalten von A, dann gilt

$$P(\pi) \circ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} z_{\pi^{-1}(1)} \\ \vdots \\ z_{\pi^{-1}(n)} \end{pmatrix} .$$

Im Fall $m = n$ und $A = E_n$ ergibt sich die Formel aus (7) *für die Permutationsmatrix $P(\pi)$.*

(9) *Seien $\rho \in \mathrm{S}_m$ und $A = (a_{\nu\mu}) \in \mathbb{R}^{n \times m}$. Dann gilt*

$$A \circ P(\rho) = (\alpha_{\nu\mu}) \,, \quad \alpha_{\nu\mu} = a_{\nu\rho(\mu)}$$

für alle $\nu = 1, \ldots, n$ und alle $\mu = 1, \ldots, m$. Folglich ist die $\rho^{-1}(\mu)$-te Spalte von $A \circ P(\rho)$ gleich der μ-ten Spalte von A. Sind $a_1, \ldots, a_m \in \mathbb{R}^n$ die Spalten von A, dann gilt

$$(a_1, \ldots, a_m) \circ P(\rho) = (a_{\rho(1)}, \ldots, a_{\rho(m)}) \,.$$

Im Fall $m = n$ und $A = E_n$ ergibt sich die definierende Gleichung aus (5) *für die Permutationsmatrix $P(\rho)$.*

(10) *Seien $\pi \in \mathrm{S}_n$ und $A = (a_{\nu\mu}) \in \mathbb{R}^{n \times n}$. Dann gilt*

$$(P(\pi))^{-1} \circ A \circ P(\pi) = (\gamma_{\nu\mu}) \,, \quad \gamma_{\nu\mu} = a_{\pi(\nu)\pi(\mu)}$$

für alle $\nu, \mu = 1, \ldots, n$. Beim Übergang von der quadratischen Matrix A zur quadratischen Matrix

$$(P(\pi))^{-1} \circ A \circ P(\pi)$$

werden die Diagonalelemente permutiert.

(11) *Es gelten:*

(11.1) $\mathrm{Perm}(n, \mathbb{R}) \cap \mathrm{L}^{\llcorner}_{\times}(n, \mathbb{R}) = \{E_n\}$.

(11.2) $\mathrm{Perm}(n, \mathbb{R}) \cap \mathrm{U}^{\urcorner}_{\times}(n, \mathbb{R}) = \{E_n\}$.

Siehe Satz und Definition 9.2, *Aussage* (4), *insbesondere* (4.5) *und* (4.6).

Beweis. Die Aussagen (1.1), (1.2), (1.3) und (1.4) sind klar. Ebenso sind die Aussagen (4.1), (4.3) und (4.5) klar. Offenbar gibt es zu jeder Permutationsmatrix $P \in \mathrm{Perm}(n, \mathbb{R})$ genau eine Permutationsmatrix $\pi \in \mathrm{S}_n$ mit $P = P(\pi)$. Die Aussagen (1.4), (4.2) und (4.4) folgen aus (4.5), (6.2) und (6.3).

Zu Aussage (6). Für π, $\sigma \in \mathrm{S}_n$ und $\mu = 1, \ldots, n$ gilt

$$(P(\pi) \circ P(\sigma))e_\mu = P(\pi)(P(\sigma)e_\mu)$$

$$
\begin{aligned}
&= P(\pi)e_{\sigma(\mu)} \\
&= e_{\pi(\sigma(\mu))} \\
&= e_{(\pi\circ\sigma)(\mu)} \\
&= P(\pi\circ\sigma)e_{\mu}\,.
\end{aligned}
$$

Damit sind (6.2) und (4.2) bewiesen. Sei $P = P(\pi) \in \mathrm{Perm}(n,\mathbb{R})$ mit $\pi \in \mathrm{S}_n$ gegeben. Dann gilt

$$
P^tP = \begin{pmatrix} e^t_{\pi(1)} \\ \vdots \\ e^t_{\pi(n)} \end{pmatrix} (e_{\pi(1)},\dots,e_{\pi(n)}) = (\langle e_{\pi(\nu)}, e_{\pi(\mu)}\rangle) = (\delta_{\nu\mu}) = E_n\,.
$$

Es folgt $P^{-1} = P^t$. Damit sind (6.3) und (4.4) bewiesen. Wir beweisen (6.4). Wegen $\pi_{i\leftrightarrow j} = \pi^{-1}_{i\leftrightarrow j}$ ist die zugehörige Permuationsmatrix $P(\pi_{i\leftrightarrow j})$ nach Aussage (6.3) symmetrisch. Es genügt, den Fall $i < j$ zu betrachten. Dann gilt

$$
\begin{aligned}
P(\pi_{i\leftrightarrow j}) &= (\dots, e_j, \dots, e_i, \dots) \\
&= (\dots, e_j, \dots, e_i, \dots)^t \\
&= \begin{pmatrix} \vdots \\ e^t_j \\ \vdots \\ e^t_i \\ \vdots \end{pmatrix} = E_{i\leftrightarrow j}\,.
\end{aligned}
$$

Damit ist auch (6.4) bewiesen.

Zu Aussage (7). Sei $\pi \in \mathrm{S}_n$. Nach Definition (5) gilt

$$
P(\pi) = (p_{\nu\mu}) = (e_{\pi(1)},\dots,e_{\pi(n)})\,.
$$

Der Eintrag $p_{\nu\mu}$ von $P(\pi) \in \mathrm{Perm}(n,\mathbb{R})$ ist genau dann gleich 1, wenn $\nu = \pi(\mu)$ gilt. Dies ist mit $\pi^{-1}(\nu) = \mu$ äquivalent. Also gilt

$$
p_{\nu\mu} = \langle e_\nu, P(\pi)e_\mu\rangle = \delta_{\nu\pi(\mu)} = \delta_{\pi^{-1}(\nu)\mu}
$$

für alle $\nu,\mu = 1,\dots,n$. In Matrixschreibweise bedeuten diese Gleichungen

$$
P(\pi) = (e_{\pi(1)},\dots,e_{\pi(n)}) = \begin{pmatrix} e^t_{\pi^{-1}(1)} \\ \vdots \\ e^t_{\pi^{-1}(n)} \end{pmatrix}.
$$

Damit ist (7) bewiesen.

Zu Aussage (8). Seien $\pi \in \mathrm{S}_n$, $P(\pi) = (p_{\nu\mu}) \in \mathrm{Perm}(n, \mathbb{R})$, $A = (a_{\lambda\mu}) \in \mathbb{R}^{n\times m}$ und $P(\pi) \circ A = (\alpha_{\nu\mu}) \in \mathbb{R}^{n\times m}$. Aus (7) folgt

$$\begin{aligned} \alpha_{\nu\mu} &= \sum_{\lambda=1}^{n} p_{\nu\lambda}\, a_{\lambda\mu} \\ &= \sum_{\lambda=1}^{n} \delta_{\pi^{-1}(\nu)\,\lambda}\, a_{\lambda\mu} \\ &= a_{\pi^{-1}(\nu)\,\mu}\,. \end{aligned}$$

für alle $\nu, \mu = 1, \ldots, n$. Sind $z_1^t, \ldots, z_n^t \in \mathbb{R}^m$ die Spalten von A, dann gilt

$$P(\pi) \circ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} z_{\pi^{-1}(1)} \\ \vdots \\ z_{\pi^{-1}(n)} \end{pmatrix}.$$

Damit ist (8) bewiesen.

Zu Aussage (9). Seien $\rho \in \mathrm{S}_m$, $P(\rho) = (p_{\lambda\mu}) \in \mathrm{Perm}(m, \mathbb{R})$, $A = (a_{\nu\lambda}) \in \mathbb{R}^{n\times m}$ und $A \circ P(\pi) = (\alpha_{\nu\mu}) \in \mathbb{R}^{n\times m}$. Aus (7) folgt

$$\begin{aligned} \alpha_{\nu\mu} &= \sum_{\lambda=1}^{m} a_{\nu\lambda}\, p_{\lambda\mu} \\ &= \sum_{\lambda=1}^{m} a_{\nu\lambda}\, \delta_{\lambda\,\rho(\mu)} \\ &= a_{\nu\,\rho(\mu)} \end{aligned}$$

für alle $\nu, \mu = 1, \ldots, n$. Sind $a_1, \ldots, a_m \in \mathbb{R}^n$ die Spalten von A, dann gilt

$$(a_1, \ldots, a_m) \circ P(\rho) = (a_{\rho(1)}, \ldots, a_{\rho(n)})\,.$$

Damit ist (9) bewiesen.

Zu Aussage (10). Für $\pi \in \mathrm{S}_n$, $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times n}$ und

$$(P(\pi))^{-1} \circ A = (\beta_{\nu\lambda}) \in \mathbb{R}^{n\times n}\,,$$

$$(P(\pi))^{-1} \circ A \circ P(\pi) = (\gamma_{\nu\mu}) \in \mathbb{R}^{n\times n}\,.$$

erhalten wir

$$\begin{aligned} \gamma_{\nu\mu} &= \sum_{\lambda=1}^{n} \beta_{\nu\lambda}\, p_{\lambda\mu} \\ &= \sum_{\lambda=1}^{n} a_{(\pi^{-1})^{-1}(\nu)\,\lambda}\, p_{\lambda\mu} \end{aligned}$$

$$= \sum_{\lambda=1}^{n} a_{(\pi^{-1})^{-1}(\nu)\,\lambda}\, \delta_{\lambda\,\pi(\mu)}$$

$$= \sum_{\lambda=1}^{n} a_{\pi(\nu)\,\lambda}\, \delta_{\lambda\,\pi(\mu)}$$

$$= a_{\pi(\nu)\,\pi(\mu)}\,.$$

Damit ist auch (10) bewiesen.

Zu Aussage (11). Es genügt, (11.1) zu beweisen. Sei

$$P \in \mathrm{Perm}(n,\mathbb{R}) \cap \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R})$$

gegeben. Dann gilt

$$P^{-1} = P^t \in \mathrm{Perm}(n,\mathbb{R}) \cap \mathrm{L}^{\llcorner}_{\times}(n,\mathbb{R}) \cap \mathrm{U}^{\urcorner}_{\times}(n,\mathbb{R}) = \{E_n\}\,.$$

Folglich gilt $P = E_n$. □

Beispiel 9.9.

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ \pi(1) & \pi(2) & \pi(3) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \in \mathrm{S}_3\,.$$

$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ \pi^{-1}(1) & \pi^{-1}(2) & \pi^{-1}(3) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathrm{S}_3\,.$$

$$\begin{aligned} P(\pi) = (e_{\pi(1)}, e_{\pi(2)}, e_{\pi(3)}) = (e_3, e_1, e_2) \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} e_2^t \\ e_3^t \\ e_1^t \end{pmatrix} \\ &= \begin{pmatrix} e_{\pi^{-1}(1)}^t \\ e_{\pi^{-1}(2)}^t \\ e_{\pi^{-1}(3)}^t \end{pmatrix} \in \mathrm{Perm}(3,\mathbb{R})\,. \end{aligned}$$

$$
\begin{aligned}
(P(\pi))^t &= \begin{pmatrix} e^t_{\pi^{-1}(1)} \\ e^t_{\pi^{-1}(2)} \\ e^t_{\pi^{-1}(3)} \end{pmatrix}^t \\
&= (e_{\pi^{-1}(1)}, e_{\pi^{-1}(2)}, e_{\pi^{-1}(3)}) \\
&= P(\pi^{-1}) \\
&= (P(\pi))^{-1} \,.
\end{aligned}
$$

$$
(P(\pi))^{-1} = (P(\pi))^t = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \,.
$$

$$
\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \,.
$$

$$
\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \,.
$$

$$
A = (a_1, a_2, a_3) = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathbb{R}^3 \,.
$$

$$
\begin{aligned}
P(\pi) \circ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \\
&= \begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{pmatrix} \\
&= \begin{pmatrix} z_2 \\ z_3 \\ z_1 \end{pmatrix} \\
&= \begin{pmatrix} z_{\pi^{-1}(1)} \\ z_{\pi^{-1}(2)} \\ z_{\pi^{-1}(3)} \end{pmatrix} \,.
\end{aligned}
$$

$$
\begin{aligned}
(a_1, a_2, a_3) \circ P(\pi) &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 3 & 1 & 2 \\ 6 & 4 & 5 \\ 9 & 7 & 8 \end{pmatrix} \\
&= (a_3, a_1, a_2) \\
&= (a_{\pi(1)}, a_{\pi(2)}, a_{\pi(3)})\,.
\end{aligned}
$$

$$
\begin{aligned}
(P(\pi))^{-1} \circ A \circ P(\pi) = (P(\pi))^t \circ A \circ P(\pi) \\
&= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 & 2 \\ 6 & 4 & 5 \\ 9 & 7 & 8 \end{pmatrix} \\
&= \begin{pmatrix} 9 & 7 & 8 \\ 3 & 1 & 2 \\ 6 & 4 & 5 \end{pmatrix}\,.
\end{aligned}
$$

$\square$

Nach Satz 9.8, Aussage (8) bewirkt eine linksseitige Multiplikation

$$A \mapsto P(\pi) \circ A$$

mit einer $(n \times n)$-Permutationsmatrix $P(\pi)$ eine Vertauschung der Zeilen einer $(n \times m)$-Matrix A. Die Formel

$$P(\pi) \circ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} z_{\pi^{-1}(1)} \\ \vdots \\ z_{\pi^{-1}(n)} \end{pmatrix}$$

bedeutet, dass die $\pi(\nu)$-te Zeile von $P(\pi) \circ A$ gleich der ν-ten Zeile von A ist. Beim Übergang

$$A \mapsto P(\pi) \circ A$$

wird die ν-te Zeile von A in die $\pi(\nu)$-te Zeile von $P(\pi) \circ A$ verschoben. Dies gilt für jeden Zeilenindex $\nu = 1, \ldots, n$.

$$\begin{array}{c||c} A & \\ \hline & Z_{\pi^{-1}(1)} \\ P(\pi) \circ A & \vdots \\ & Z_{\pi^{-1}(n)} \end{array} \qquad \begin{array}{c||c} E_n & \\ \hline & Z_{\pi^{-1}(1)} \\ P(\pi) & \vdots \\ & Z_{\pi^{-1}(n)} \end{array}$$

Wir weisen auf die entsprechende Erörterung der Elementarmatrizen im Anschluss an Satz und Definition 6.2 hin.

Beispiel 9.10 (Fortsetzung von 9.9)**.** Wir betrachten noch einmal

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad P(\pi) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Die Übergänge $E_3 \mapsto P(\pi)$ und $A \mapsto P(\pi) \circ A$ beschreiben wir schematisch, wie folgt.

$$\begin{array}{ccc||c} 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ \hline 0 & 1 & 0 & Z_2 \\ 0 & 0 & 1 & Z_3 \\ 1 & 0 & 0 & Z_1 \end{array} \qquad \begin{array}{ccc||c} 1 & 2 & 3 & \\ 4 & 5 & 6 & \\ 7 & 8 & 9 & \\ \hline 4 & 5 & 6 & Z_2 \\ 7 & 8 & 9 & Z_3 \\ 1 & 2 & 3 & Z_1 \end{array}$$

□

10 Strikte Bruhat-Zerlegung

Satz 10.1 (Bruhat-Zerlegung)**.** *Seien $n \in \mathbb{N}$ und $A \in \mathrm{GL}(n, \mathbb{R})$ beliebig gegeben. Dann gibt es Matrizen $L \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$, $D \in \mathrm{D}^{\times}(n, \mathbb{R})$, $P \in \mathrm{Perm}(n, \mathbb{R})$ und $U \in \mathrm{U}_1^{\urcorner}(n, \mathbb{R})$ mit*

$$A = L \circ D \circ P \circ U\,.$$

Eine solche Produktzerlegung heißt eine Bruhat-Zerlegung *der invertierbaren Matrix A. Es gelten die folgenden Einzigkeitsaussagen* (1) *und* (2)*.*

(1) *Die Matrizen D und P sind durch A eindeutig bestimmt.*

(2) *Die Matrix $U \in \mathrm{U}_1^{\urcorner}(n, \mathbb{R})$ kann so gewählt werden, dass die Bedingung*

$$P \circ U \circ P^{-1} \in \mathrm{U}_1^{\urcorner}(n, \mathbb{R})$$

erfüllt ist. Wenn diese Bedingung erfüllt ist, dann sind L, D, P und U eindeutig durch A bestimmt. In diesem Fall nennen wir die Bruhat-Zerlegung die strikte *Bruhat-Zerlegung von A.*

Beweis. Offenbar gilt

$$(\alpha) = (1) \circ (\alpha) \circ (1) \circ (1)$$

für alle $\alpha \in \mathbb{R}^{\times}$. Die Aussagen des Satzes sind daher im Fall $n = 1$ trivialerweise erfüllt. Wir betrachten daher den Fall $n \geq 2$.

Wir beweisen zuerst die Einzigkeitsaussagen. Seien

$$A = L \circ D \circ P \circ U = l \circ d \circ p \circ u$$

zwei Bruhat-Zerlegungen von A. Dann ist

$$B = d^{-1} \circ l^{-1} \circ L \circ D = p \circ u \circ U^{-1} \circ P^{-1}$$

eine invertierbare untere Dreiecksmatrix. Es gilt

$$u \circ U^{-1} = E_n + N\,,$$

wobei N eine obere Dreiecksmatrix ist, deren sämtliche Diagonalelemente verschwinden. Es gibt daher keine Position, an der E_n und N zugleich nichtverschwindende Einträge besitzen. Diese Eigenschaft überträgt sich nach (8) und (9) aus 9.8 auf die Summanden $p \circ P^{-1} = p \circ E_n \circ P^{-1}$ und $p \circ N \circ P^{-1}$ der Matrix B. Weil B eine untere Dreiecksmatrix ist, kann kein Pivotelement von $p \circ P^{-1}$ oberhalb der Diagonale stehen. Also gilt $p \circ P^{-1} = E_n$. Es folgt $p = P$. Nach (10) in 9.8 bedeutet dies, dass alle Diagonalelemente von

$$B = p \circ u \circ U^{-1} \circ P^{-1} = P \circ u \circ U^{-1} \circ P^{-1}$$

gleich 1 sind. Weil $l^{-1} \circ L$ dieselbe Eigenschaft besitzt, folgt $d = D$.

Wir nehmen zusätzlich an, dass die obigen Bruhat-Zerlegungen strikt sind. Dann folgt aus $p = P$, dass

$$B = (p \circ u \circ p^{-1}) \circ (P \circ U^{-1} \circ P^{-1}) \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R}) \cap \mathrm{U}_1^{\urcorner}(n, \mathbb{R}) = \{E_n\}$$

gilt. Also gilt $B = E_n$. Wegen $d = D$ und $p = P$ folgen nun $l = L$ und $u = U$. Damit sind die Einzigkeitsaussagen (1) und (2) bewiesen.

Wir zeigen durch Angabe eines Konstruktionsverfahrens, dass eine beliebige invertierbare Matrix A eine strikte Bruhat-Zerlegung besitzt.

Weil A invertierbar ist, enthält die erste Zeile von A einen nicht-verschwindenden Eintrag. Sei $d_1 \in \mathbb{R}^\times$ der nicht-verschwindende Eintrag der ersten Zeile von A mit dem kleinsten Spaltenindex $\sigma(1)$. Durch $(n-1)$ elementare Zeilenumformungen der Form

$$Z_i - l_{i1} Z_1$$

mit $i = 2, \ldots, n$ und $l_{i1} \in \mathbb{R}$ können die Einträge unterhalb der Position $(1, \sigma(1))$ annulliert werden. Die Matrix

$$A^{(1)} = \left(\prod_{i=2}^{n} E_{i;-l_{i1};1} \right) \circ A \in \mathbb{R}^{n \times n}$$

ist invertierbar. Es gilt

$$L_1 = \prod_{i=2}^{n} E_{i;l_{i1};1} \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})\,.$$

Die $(n-1)$ Faktoren $E_{i;l_{i1};1}$ mit $i = 2, \ldots, n$ der unteren Dreiecksmatrix L_1 sind paarweise vertauschbar.

Weil $A^{(1)}$ invertierbar ist, enthält die zweite Zeile von $A^{(1)}$ einen nicht-verschwindenden Eintrag. Sei $d_2 \in \mathbb{R}^\times$ der nicht-verschwindende Eintrag der zweiten Zeile von $A^{(1)}$ mit dem kleinsten Spaltenindex $\sigma(2)$. Im Fall $n > 2$ können durch weitere $(n-2)$ elementare Zeilenumformungen der Form

$$Z_i - l_{i2} Z_2$$

mit $i = 3, \ldots, n$ und $l_{i2} \in \mathbb{R}$ die Einträge unterhalb der Position $(2, \sigma(2))$ annulliert werden. Die Matrix

$$A^{(2)} = \left(\prod_{i=3}^{n} E_{i;-l_{i2};2} \right) \circ A^{(1)} \in \mathbb{R}^{n \times n}$$

ist invertierbar. Es gilt

$$L_2 = \prod_{i=3}^{n} E_{i;l_{i2};2} \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})\,.$$

Die $(n-2)$ Faktoren $E_{i;l_{i2};2}$ mit $i = 3, \ldots, n$ der unteren Dreiecksmatrix L_2 sind paarweise vertauschbar.

Weil $A^{(2)}$ invertierbar ist, enthält die dritte Zeile von $A^{(2)}$ einen nicht-verschwindenden Eintrag. Sei $d_3 \in \mathbb{R}^\times$ der nicht-verschwindende Eintrag der dritten Zeile von $A^{(2)}$ mit dem kleinsten Spaltenindex $\sigma(3)$.

Durch Wiederholung der Konstruktion erhalten wir eine Permutation

$$\sigma = \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix} \in \mathrm{S}_n \,,$$

die zugehörige Permutationsmatrix

$$P(\sigma) = (e_{\sigma(1)}, \dots, e_{\sigma(n)}) \in \mathrm{Perm}(n, \mathbb{R}) \,,$$

eine invertierbare Matrix

$$A^{(n-1)} = (\alpha_{\nu\mu}) \in \mathrm{GL}(n, \mathbb{R}) \,,$$

deren Einträge links und unterhalb der Positionen $(i, \sigma(i))$ verschwinden, eine Diagonalmatrix

$$D = \mathrm{diag}(d_1, \dots, d_n) = \mathrm{diag}(\alpha_{1\sigma(1)}, \dots, \alpha_{n\sigma(n)}) \in \mathrm{D}^\times(n, \mathbb{R})$$

sowie untere Dreiecksmatrizen $L_1, \dots, L_{n-1} \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$ mit

$$A = \left(\prod_{i=1}^{n-1} L_i \right) \circ A^{(n-1)} = L_1 \circ \ldots \circ L_{n-1} \circ A^{(n-1)} \,.$$

Die Matrizen $L_1, \dots, L_{n-1}$ sind im Allgemeinen nicht vertauschbar. Es gilt

$$L = L_1 \circ \ldots \circ L_{n-1} \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R}) \,.$$

Der Eintrag von L unterhalb der Hauptdiagonalen an der Position (ν, μ) mit $\nu > \mu$ ist gleich der oben konstruierten reellen Zahl $l_{\nu\mu}$. Siehe Satz 9.4 und die Beispiele 9.5 und 10.4. Wir setzen

$$P = (P(\sigma))^{-1} = (P(\sigma))^t \in \mathrm{Perm}(n, \mathbb{R}) \,.$$

Nach (8) in 9.8 ist die $\sigma(i)$-te Zeile von

$$U = P(\sigma) \circ D^{-1} \circ A^{(n-1)} = P(\sigma) \circ \mathrm{diag}(d_1^{-1}, \dots, d_n^{-1}) \circ A^{(n-1)}$$

gleich der i-ten Zeile von $D^{-1} \circ A^{(n-1)}$. Also gilt

$$U \in \mathrm{U}_1^{\urcorner}(n, \mathbb{R}) \,.$$

Nach Konstruktion ist

$$A = L \circ D \circ P \circ U$$

eine Bruhat-Zerlegung von A. Wir zeigen abschließend, dass diese Zerlegung sogar eine strikte Bruhat-Zerlegung ist. Nach (9) in 9.8 ist die $\sigma^{-1}(i)$-te Spalte von

$$V = P \circ U \circ P^{-1} = (P(\sigma))^{-1} \circ U \circ P(\sigma) = D^{-1} \circ A^{(n-1)} \circ P(\sigma)$$

gleich der i-ten Spalte von $D^{-1} \circ A^{(n-1)}$. Die Komponenten der i-ten Spalte von $A^{(n-1)}$ verschwinden unterhalb der Position $(\sigma^{-1}(i), i)$. Also verschwinden alle Einträge von V unterhalb der Hauptdiagonalen. Nach (10) in 9.8 sind alle Diagonalelemente von V gleich 1. Also ist auch die Bedingung $P \circ U \circ P^{-1} \in \mathrm{U}_1^{\urcorner}(n, \mathbb{R})$ erfüllt. $\square$

Beispiel 10.2. Gegeben sei die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix}.$$

Wir berechnen die strikte Bruhat-Zerlegung von A mit dem Verfahren aus dem Beweis des Satzes 10.1. Anschießend vereinfachen wir das Rechenschema.

$$\begin{array}{rrr|l}
0 & 0 & -2 & \\
-1 & -3 & -10 & \\
-4 & -9 & 2 & \\
\hline
0 & 0 & -2 & Z_1 \\
-1 & -3 & 0 & Z_2 - (+5)Z_1 \\
-4 & -9 & 0 & Z_3 - (-1)Z_1 \\
\hline
0 & 0 & -2 & Z_1 \\
-1 & -3 & 0 & Z_2 \\
0 & 3 & 0 & Z_3 - (+4)Z_2 \\
\hline\hline
0 & 0 & 1 & (-2)^{-1}Z_1 \\
1 & 3 & 0 & (-1)^{-1}Z_2 \\
0 & 1 & 0 & (+3)^{-1}Z_3
\end{array}$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -1 & 4 & 1 \end{pmatrix}, \quad D = \operatorname{diag}(-2,-1,3) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

$$\sigma(1) = 3\,, \quad \sigma(2) = 1\,, \quad \sigma(3) = 2\,.$$

$$P(\sigma) = (e_{\sigma(1)}, e_{\sigma(2)}, e_{\sigma(3)}) = (e_3, e_1, e_2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

$$P = (P(\sigma))^{-1} = (P(\sigma))^t = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

$$U = P(\sigma) \circ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nun vereinfachen wir das Rechenschema. Nach Konstruktion steht das Pivotelement der i-ten Zeile von

$$D^{-1} \circ A^{(2)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

an der Position $(i, \sigma(i))$. Durch linksseitige Multiplikation mit der Permutationsmatrix $P(\sigma)$ wird die i-te Zeile der Matrix $D^{-1} \circ A^{(2)}$ in die $\sigma(i)$-te Zeile der oberen Dreiecksmatrix

$$U = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

verschoben. Die linksseitige Multiplikation mit der Permutationsmatrix

$$P = (P(\sigma))^{-1} = (P(\sigma))^t = \begin{pmatrix} e^t_{\sigma(1)} \\ e^t_{\sigma(2)} \\ e^t_{\sigma(3)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

macht die genannten Zeilenverschiebungen rückgängig. Offenbar besetzen die Pivotelemente der beiden Matrizen $D^{-1} \circ A^{(2)}$ und P dieselben Positionen. Daher kann die Permutationsmatrix P an der pivotnormierten Matrix $D^{-1} \circ A^{(2)}$ abgelesen werden: *P entsteht aus $D^{-1} \circ A^{(2)}$ durch Annullierung aller Einträge, die keine Pivotelemente sind.* Siehe Satz 10.5.

$$\begin{array}{rrr|l}
0 & 0 & -2 & \\
-1 & -3 & -10 & \\
-4 & -9 & 2 & \\
\hline
0 & 0 & -2 & Z_1 \\
-1 & -3 & 0 & Z_2 - (+5)Z_1 \\
-4 & -9 & 0 & Z_3 - (-1)Z_1 \\
\hline
0 & 0 & -2 & Z_1 \\
-1 & -3 & 0 & Z_2 \\
0 & 3 & 0 & Z_3 - (+4)Z_2 \\
\hline\hline
0 & 0 & 1 & (-2)^{-1}Z_1 \\
1 & 3 & 0 & (-1)^{-1}Z_2 \\
0 & 1 & 0 & (+3)^{-1}Z_3 \\
\hline\hline
0 & 0 & 1 & \\
1 & 0 & 0 & \\
0 & 1 & 0 &
\end{array}$$

Wir fassen die wichtigsten Schritte der Rechnung noch einmal zusammen. Die Übertragung auf den allgemeinen Fall ist offenkundig. Siehe das folgende Beispiel 10.3.

- Die elementaren Zeilenumformungen, die auf die Matrizen $A^{(1)}$ und $A^{(2)}$ führen, ergeben die Einträge unterhalb der Hauptdiagonale der Matrix L.
- Die Pivotelemente von $A^{(2)}$ sind die Diagonalelemente von D.
- **Normieren** der Pivotelemente der Matrix $A^{(2)}$ mit elementaren Zeilenumformungen liefert die Matrix $D^{-1} \circ A^{(2)}$.
- Umordnen der Zeilen der Matrix $D^{-1} \circ A^{(2)}$ liefert die Matrix U.
- **Nullsetzen der Nicht-Pivotelemente** der Matrix $D^{-1} \circ A^{(2)}$ liefert die Permutationsmatrix P.

In 12.3 definieren wir die Determinante einer invertierbaren Matrix mit Hilfe ihrer strikten Bruhat-Zerlegung. Die Determinante der Matrix A berechnen wir in Beispiel 12.4.

$\square$

Beispiel 10.3.

$$
\begin{array}{rrrr|l}
0 & -3 & -6 & -3 & \\
10 & 9 & 98 & 39 & \\
-5 & -3 & -46 & -14 & \\
0 & -3 & -8 & 7 & \\
\hline
0 & -3 & -6 & -3 & Z_1 \\
10 & 0 & 80 & 30 & Z_2 - (-3)Z_1 \\
-5 & 0 & -40 & -11 & Z_3 - (+1)Z_1 \\
0 & 0 & -2 & 10 & Z_4 - (+1)Z_1 \\
\hline
0 & -3 & -6 & -3 & Z_1 \\
10 & 0 & 80 & 30 & Z_2 \\
0 & 0 & 0 & 4 & Z_3 - (-\frac{1}{2})Z_2 \\
0 & 0 & -2 & 10 & Z_4 - 0 \cdot Z_2 \\
\hline
0 & -3 & -6 & -3 & Z_1 \\
10 & 0 & 80 & 30 & Z_2 \\
0 & 0 & 0 & 4 & Z_3 \\
0 & 0 & -2 & 0 & Z_4 - (+\frac{5}{2})Z_3 \\
\hline\hline
0 & 1 & 2 & 1 & (-3)^{-1}Z_1 \\
1 & 0 & 8 & 3 & (+10)^{-1}Z_2 \\
0 & 0 & 0 & 1 & (+4)^{-1}Z_3 \\
0 & 0 & 1 & 0 & (-2)^{-1}Z_4 \\
\hline\hline
0 & 1 & 0 & 0 & \\
1 & 0 & 0 & 0 & \\
0 & 0 & 0 & 1 & \\
0 & 0 & 1 & 0 &
\end{array}
$$

$$
L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 1 & -\frac{1}{2} & 1 & 0 \\ 1 & 0 & \frac{5}{2} & 1 \end{pmatrix}, \quad D = \operatorname{diag}(-3, 10, 4, -2)\,.
$$

$$
P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 0 & 8 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
$$

$\square$

Beispiel 10.4. Wir betrachten die Konstruktion der unteren Dreiecksmatrix

$$L = L_1 \circ \ldots \circ L_{n-1} \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$$

in den Fällen $n = 2, 3, 4$ im Detail. Im allgemeine Fall folgt durch wiederholte Anwendung von Satz 9.4. Wir verwenden die Bezeichnungen aus dem Beweis des Satzes 10.1 über die Bruhat-Zerlegung.

Sei $n = 2$. Dann gilt

$$L = L_1 = \begin{pmatrix} 1 & 0 \\ l_{21} & 1 \end{pmatrix} = E^{(2)}_{2;l_{21};1} .$$

Sei $n = 3$. Dann gilt

$$L = L_1 \circ L_2 = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} .$$

Beachten Sie bitte die Reihenfolge der Matrizen. Die unteren Dreiecksmatrizen L_1 und L_2 sind im Allgemeinen nicht vertauschbar. Es gelten

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E^{(3)}_{3;l_{31};1} \circ E^{(3)}_{2;l_{21};1} ,$$

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{32} & 1 \end{pmatrix} = E^{(3)}_{3;l_{32};2} .$$

Die Matrizen $E^{(3)}_{3;l_{31};1}$ und $E^{(3)}_{2;l_{21};1}$ sind vertauschbar.

Sei $n = 4$. Dann gilt

$$\begin{aligned} L &= L_1 \circ L_2 \circ L_3 \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{42} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & l_{43} & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} . \end{aligned}$$

Beachten Sie bitte die Reihenfolge der Matrizen. Die unteren Dreiecksmatrizen L_1, L_2 und L_3 sind im Allgemeinen nicht vertauschbar. Wir schreiben L_1, L_2 und L_3 als Produkte von Elementarmatrizen. Es gilt

$$L_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ l_{41} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ l_{31} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= E^{(4)}_{4;l_{41},1} \circ E^{(4)}_{3;l_{31},1} \circ E^{(4)}_{2;l_{21};1}\,.$$

Die drei Matrizen $E^{(4)}_{4;l_{41},1}$, $E^{(4)}_{3;l_{31},1}$, $E^{(4)}_{2;l_{21};1}$ sind paarweise vertauschbar. Es gilt

$$L_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & l_{42} & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & l_{42} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = E^{(4)}_{4;l_{42};2} \circ E^{(4)}_{3;l_{32};2}\,.$$

Die beiden Matrizen $E^{(4)}_{4;l_{42};2}$ und $E^{(4)}_{3;l_{32};2}$ sind vertauschbar. Es gilt

$$L_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & l_{43} & 1 \end{pmatrix} = E^{(4)}_{4;l_{43};3}\,.$$

□

Wir stellen die Vereinfachungen des Rechenschemas zusammen, die wir in Beispiel 10.2 erörtet haben. Wir verwenden dabei die dortigen Definitionen sowie die Bezeichnungen aus dem Beweis des Satzes 10.1.

Satz 10.5. *Seien $n \in \mathbb{N}$ mit $n \geq 2$ und $A \in \mathrm{GL}(n, \mathbb{R})$ gegeben. Sei*

$$A = L \circ D \circ P \circ U$$

die strikte Bruhat-Zerlegung von A. Dann entsteht die Permutationsmatrix P aus der pivotnormierten Matrix

$$D^{-1} \circ A^{(n-1)} = D^{-1} \circ L^{-1} \circ A$$

durch Annullierung aller Einträge, die keine Pivotelemente sind.

Beweis. Die i-te Zeile von U ist nach Konstruktion gleich der $\sigma(i)$-ten Zeile von

$$D^{-1}A^{(n-1)} = D^{-1}L^{-1}A = PU\,.$$

Die Pivotelemente von $D^{-1}A^{(n-1)}$ stehen an den Positionen $(i, \sigma(i))$. Wir betrachten die Permuationsmatrix

$$P = P(\sigma^{-1}) = (p_{\nu\mu})\,.$$

Nach Satz 9.8 Aussage (7) gilt

$$p_{i\sigma(i)} = \delta_{(\sigma^{-1})^{-1}(i)\sigma(i)} = \delta_{\sigma(i)\sigma(i)} = 1$$

für alle $i = 1, \ldots, n$. Folglich stehen die Pivotelemente von $D^{-1}A^{(n-1)}$ und P an denselben Positionen. □

Beispiel 10.6. Sei $n \in \mathbb{N}$. Wir betrachten die strikten Bruhat-Zerlegungen spezieller invertierbarer Matrizen.

(1) $(\forall L \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})) : L = L \circ E_n \circ E_n \circ E_n$.

(2) $(\forall D \in \mathrm{D}^{\times}(n, \mathbb{R})) : D = E_n \circ D \circ E_n \circ E_n$.

(3) $(\forall P \in \mathrm{Perm}(n, \mathbb{R})) : P = E_n \circ E_n \circ P \circ E_n$.

(4) $(\forall U \in \mathrm{U}_1^{\urcorner}(n, \mathbb{R})) : U = E_n \circ E_n \circ E_n \circ U$.

□

Beispiel 10.7. Wir geben die strikte Bruhat-Zerlegung im Fall $n = 2$ explizit an. Sei

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

eine invertierbare reelle (2×2)-Matrix. Dann gilt $\alpha\delta - \gamma\beta \neq 0$.

(1) Sei $\alpha \in \mathbb{R}^{\times}$. In diesem Fall gilt $\delta - \gamma\alpha^{-1}\beta \neq 0$. Das Schema

$$\begin{array}{cc|l}
\alpha & \beta & \\
\gamma & \delta & \\
\hline
\alpha & \beta & Z_1 \\
0 & \delta - \gamma\alpha^{-1}\beta & Z_2 - \gamma\alpha^{-1}Z_1 \\
\hline\hline
1 & \alpha^{-1}\beta & \alpha^{-1}Z_1 \\
0 & 1 & (\delta - \gamma\alpha^{-1}\beta)^{-1}Z_2 \\
\hline\hline
1 & 0 & \\
0 & 1 &
\end{array}$$

liefert

$$A^{(1)} = \begin{pmatrix} \alpha & \beta \\ 0 & \delta - \gamma\alpha^{-1}\beta \end{pmatrix}.$$

und die strikte Bruhat-Zerlegung

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \gamma\alpha^{-1} & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \delta - \gamma\alpha^{-1}\beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha^{-1}\beta \\ 0 & 1 \end{pmatrix}.$$

(2) Sei $\alpha = 0$. In diesem Fall gilt $\gamma\beta \neq 0$. Das Schema

$$\begin{array}{cc|l} 0 & \beta & \\ \gamma & \delta & \\ \hline 0 & \beta & Z_1 \\ \gamma & 0 & Z_2 - \delta\beta^{-1}Z_1 \\ \hline\hline 0 & 1 & \beta^{-1}Z_1 \\ 1 & 0 & \gamma^{-1}Z_2 \\ \hline\hline 0 & 1 & \\ 1 & 0 & \end{array}$$

liefert

$$A^{(1)} = \begin{pmatrix} 0 & \beta \\ \gamma & 0 \end{pmatrix}.$$

und die strikte Bruhat-Zerlegung

$$A = \begin{pmatrix} 0 & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \delta\beta^{-1} & 1 \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$\square$

11 LU-Zerlegung

Nach Satz 10.1 sind die Diagonalmatrix D und die Permutationsmatrix P in der Bruhat-Zerlegung $A = LDPU$ einer invertierbaren Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ eindeutig bestimmt. Falls $P = E_n$ gilt, ist die Bruhat-Zerlegung sogar strikt und die Matrix

$$DPU = DE_nU = DU$$

ist eine invertierbare obere Dreiecksmatrix.

Definition 11.1. *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$ beliebig gegeben. Eine Zerlegung*

$$A = L \circ U$$

*mit $L \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$ und $U \in \mathrm{U}^{\urcorner}(n, \mathbb{R})$ heißt eine LU-*Zerlegung *der Matrix A.*

Satz 11.2. *Eine invertierbare Matrix $A = (a_{\nu\mu})_{\nu,\mu=1,\ldots,n} \in \mathbb{R}^{n \times n}$ besitzt genau dann eine LU-Zerlegung, wenn für alle $k = 1, \ldots, n-1$ die linken oberen $(k \times k)$-Teilmatrizen*

$$A_{1:k,1:k} = (a_{\nu\mu})_{\nu,\mu=1,\ldots,k}$$

ebenfalls invertierbar sind. Die LU-Zerlegung einer invertierbaren Matrix ist eindeutig bestimmt.

Der Algorithmus zur Bestimmung der strikten Bruhat-Zerlegung gibt Auskunft, ob eine LU-Zerlegung möglich ist. Wenn eine LU-Zerlegung möglich ist, liefert der Algorithumus zur Bestimmung der strikten Bruhat-Zerlegung gleichzeitig ein Verfahren zur Bestimmung der LU-Zerlegung.

Beispiel 11.3. Die invertierbare Matrix

$$A = \begin{pmatrix} 1 & -2 & 4 \\ -2 & 3 & -2 \\ 3 & -4 & -1 \end{pmatrix}$$

besitzt die LU-Zerlegung

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & 4 \\ 0 & -1 & 6 \\ 0 & 0 & -1 \end{pmatrix}.$$

Die Rechnung ist im folgenden Schema zusammengefasst.

$$
\begin{array}{rrr|r}
1 & -2 & 4 & \\
-2 & 3 & -2 & \\
3 & -4 & -1 & \\
\hline
1 & -2 & 4 & \\
0 & -1 & 6 & Z_2 - (-2) \cdot Z_1 \\
0 & 2 & -13 & Z_3 - (+3) \cdot Z_1 \\
\hline
1 & -2 & 4 & \\
0 & -1 & 6 & \\
0 & 0 & -1 & Z_3 - (-2) \cdot Z_2
\end{array}
$$

Abschließend sei darauf hingewiesen, dass die erste Zeile der Matrix A mit der ersten Zeile der Matrix U übereinstimmt. □

12 Determinante. Bruhat. Laplace. Leibniz

Wir verwenden die strikte Bruhat-Zerlegung, um die Determinante invertierbarer Matrizen zu definieren. Der Algorithmus zur Bestimmung der strikten Bruhat-Zerlegung liefert gleichzeitig ein Verfahren zur Bestimmung der Determinante mit Hilfe von elementaren Zeilenumformungen. Zunächst erklären wir das Signum einer Permutation.

Satz und Definition 12.1. *Sei $n \in \mathbb{N}$. Das* Signum $\operatorname{sgn}(\pi)$ *einer Permutation $\pi \in S_n$ wird durch*

$$\operatorname{sgn}(\pi) = \prod_{1 \le \nu < \mu \le n} \frac{\pi(\mu) - \pi(\nu)}{\mu - \nu}$$

definiert. Es gelten:

(1) $(\forall \pi \in S_n) : \operatorname{sgn}(\pi) \in \{-1, 1\}$.

(2) $\operatorname{sgn}(\mathrm{id}_{\{1,\dots,n\}}) = 1$.

(3) *Für jede Transposition $\pi \in S_n$ gilt* $\operatorname{sgn}(\pi) = -1$.

(4) $(\forall \pi, \sigma \in S_n) : \operatorname{sgn}(\pi \circ \sigma) = \operatorname{sgn}(\pi) \cdot \operatorname{sgn}(\sigma)$.

(5) $(\forall \pi \in S_n) : \operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$.

Beispiel 12.2. Für die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ \sigma(1) & \sigma(2) & \sigma(3) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \in \mathrm{S}_3$$

erhalten wir

$$\begin{aligned} \operatorname{sgn}(\sigma) &= \prod_{1 \le \nu < \mu \le n} \frac{\sigma(\mu) - \sigma(\nu)}{\mu - \nu} \\ &= \frac{\sigma(2) - \sigma(1)}{2-1} \cdot \frac{\sigma(3) - \sigma(1)}{3-1} \cdot \frac{\sigma(3) - \sigma(2)}{3-2} \\ &= \frac{(-2)}{1} \cdot \frac{(-1)}{2} \cdot \frac{1}{1} = 1\,. \end{aligned}$$

In Beispiel 12.4 berechnen wir das Signum der Permutation $\pi = \sigma^{-1} \in \mathrm{S}_3$ explit. Nach Rechenregel (5) in 12.1 gilt $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma) = 1$. □

Definition 12.3 (Determinante)**.** *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$ beliebig gegeben.*

(1) *Wir setzen* $\det(A) = 0$*, wenn $A \notin \mathrm{GL}(n, \mathbb{R})$.*

(2) *Sei* $A \in \mathrm{GL}(n, \mathbb{R})$ *und* $A = L \circ D \circ P \circ U$ *mit* $D = \mathrm{diag}(d_1, \ldots, d_n)$ *mit* $d_1, \ldots, d_n \in \mathbb{R}^\times$ *und* $P = P(\pi)$ *mit* $\pi \in \mathrm{S}_n$ *irgendeine Bruhat-Zerlegung von* A*. Wir setzen*

$$\det(A) = \mathrm{sgn}(\pi) \cdot \prod_{i=1}^{n} d_i \in \mathbb{R}^\times .$$

Nach den Sätzen 10.1 *und* 9.2 *sind* $d_1, \ldots, d_n \in \mathbb{R}^\times$ *und* $\pi \in \mathrm{S}_n$ *durch die invertierbare Matrix* A *eindeutig bestimmt.*

Das Verfahren zur Bestimmung der strikten Bruhat-Zerlegung liefert ein Verfahren zur Berechnung der Determinante einer Matrix $A \in \mathbb{R}^{n \times n}$. Das Verfahren aus dem Beweis des Satzes 10.1 bricht ab, sobald eine der Matrizen A oder $A^{(k)}$ mit $k \in \{1, \ldots, n-1\}$ eine Nullzeile enthält. In diesem Fall gilt $\det(A) = 0$. Wenn das Verfahren nicht abbricht, dann kann $\det(A)$ mit Hilfe der strikten Bruhat-Zerlegung entsprechend Teil (2) der Definition 12.3 berechnet werden.

Beispiel 12.4. Sei $A \in \mathbb{R}^{3 \times 3}$ die Matrix mit

$$A = \begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix} .$$

Nach den Ausführungen zu Beispiel 10.2 besitzt A die strikte Bruhat-Zerlegung

$$A = LDPU = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -1 & 4 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Die Determinante von A wird durch die Diagonalmatrix D und die Permutationsmatrix $P = P(\pi)$ bestimmt. Es gelten

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathrm{S}_3 , \quad \mathrm{sgn}(\pi) = \frac{3-2}{2-1} \cdot \frac{1-2}{3-1} \cdot \frac{1-3}{3-2} = 1 .$$

Siehe Beispiel 12.2. Die Auswertung der definierenden Formel in 12.3 ergibt

$$\det(A) = 1 \cdot \{(-2) \cdot (-1) \cdot 3\} = 6 .$$

Nach Satz 12.5 kann $\mathrm{sgn}(\pi)$ direkt an P abgelesen werden. Die Matrix P geht durch zwei Zeilenvertauschungen aus der Einheitsmatrix E_3 hervor. Es folgt

$$\mathrm{sgn}(\pi) = \det(P(\pi)) = (-1)^2 = 1 .$$

$\square$

In Satz 12.5 bestimmen wir die Determinante der Faktoren, die in den Bruhat-Zerlegungen auftreten, und das Verhalten der Determinante unter Zeilen- und Spaltenumformungen. In Satz 12.8 folgern wir aus den Ergebnissen des Satzes 12.5 die Multiplikativität der Determinante.

Der Satz 12.5 ermöglicht die Berechnung von $\det(A)$, indem die gegebene Matrix $A \in \mathbb{R}^{n\times n}$ durch elementare Zeilen- oder Spaltenumformungen auf Dreiecksform gebracht wird. Siehe Beispiel 12.7.

Satz 12.5 (Eigenschaften der Determinante)**.** *Seien $n \in \mathbb{N}$ und A, $B \in \mathbb{R}^{n\times n}$ beliebig gegeben.*

(1) $\det(A) \neq 0 \Leftrightarrow A \in \mathrm{GL}(n, \mathbb{R})$.

(2) $\det(E_n) = 1$.

(3) $(\forall \pi \in \mathrm{S}_n):\ \det(P(\pi)) = \mathrm{sgn}(\pi)\,.$

(4) *Wenn $A = (a_{\nu\mu})$ eine obere oder untere Dreiecksmatrix ist, dann gilt*

$$\det(A) = \prod_{\nu=1}^{n} a_{\nu\nu} = a_{11} \cdot \ldots \cdot a_{nn}\,.$$

(5) $\det(A^t) = \det(A)$.

(6) *Es gilt* $\det(B) = -\det(A)$*, wenn B aus A durch Vertauschung zweier verschiedener Zeilen entsteht.*

(7) *Es gilt* $\det(B) = \det(A)$*, wenn B aus A durch Addition einer Zeile zu einer anderen Zeile entsteht.*

(8) *Es gilt* $\det(B) = \alpha \det(A)$*, wenn B aus A durch Multiplikation einer Zeile mit $\alpha \in \mathbb{R}$ entsteht.*

(9) *Es gilt* $\det(B) = \det(A)$*, wenn B aus A dadurch entsteht, dass zu einer Zeile eine Linearkombination aus anderen Zeilen addiert wird.*

(10) *Die Determinante ist* in jeder Zeile linear. *Damit ist die folgende Eigenschaft gemeint. Seien $k \in \{1, \ldots, n\}$ und $z^t \in \mathbb{R}^n$ beliebig gewählt. Die Matrix $A'_k(z) \in \mathbb{R}^{n\times n}$ entstehe aus A, indem die k-te Zeile durch z ersetzt wird. Dann gilt*

$$\det(A'_k(\alpha x + \beta y)) = \alpha \det(A'_k(x)) + \beta \det(A'_k(y))$$

für alle α, $\beta \in \mathbb{R}$ und alle x^t, $y^t \in \mathbb{R}^n$.

Wegen (5) *gelten die zu* (6) *bis* (10) *analogen Aussagen, wenn die Spalten der Matrizen anstelle der Zeilen betrachtet und dabei die Zeilenoperationen durch die entsprechenden Spaltenoperationen ersetzt werden.*

Beweis. Für $n = 1$ gelten die Aussagen des Satzes trivialerweise. Im Folgenden sei $n \geq 2$. Die Eigenschaften (1), (2), (3) sind klar.

Nachweis von (4). Sei $A \in \mathrm{L}^{\llcorner}_{\times}(n, \mathbb{R})$ eine invertierbare untere Dreiecksmatrix. Dann besitzt A eine Bruhat-Zerlegung $A = LD = LDE_nE_n$ mit

$$d_\nu = a_{\nu\nu} \neq 0$$

für alle $\nu = 1, \ldots, n$. Mit (2) folgt

$$\det(A) = \prod_{\nu=1}^{n} d_\nu = \prod_{\nu=1}^{n} a_{\nu\nu} \,.$$

Nun sei $A \in \mathrm{U}^{\urcorner}_{\times}(n, \mathbb{R})$ eine invertierbare obere Dreiecksmatrix. Dann besitzt A eine Bruhat-Zerlegung $A = DU = E_nDE_nU$ mit

$$d_\nu = a_{\nu\nu} \neq 0$$

für alle $\nu = 1, \ldots, n$. Wieder ergibt sich

$$\det(A) = \prod_{\nu=1}^{n} d_\nu = \prod_{\nu=1}^{n} a_{\nu\nu} \,.$$

Nachweis von (5). Sei $A = LDPU$ eine Bruhat-Zerlegung von $A \in \mathrm{GL}(n, \mathbb{R})$. Dann ist

$$A^t = U^tP^tD^tL^t = U^t(P^{-1}DP)P^{-1}L^t$$

eine Bruhat-Zerlegung von A^t. Nach Aussage (10) des Satzes 9.8 ist

$$P^{-1}DP = \mathrm{diag}(d'_1, \ldots, d'_n)$$

eine Diagonalmatrix, die sich aus D durch Vertauschen der Hauptdiagonalelemente ergibt. Dabei verwenden wir

$$P = P(\pi)\,, \quad P^t = P^{-1} = P(\pi^{-1})\,, \quad \mathrm{sgn}(\pi) = \mathrm{sgn}(\pi^{-1})\,.$$

Schließlich erhalten wir

$$\det(A^t) = \mathrm{sgn}(\pi^{-1}) \cdot \prod_{\nu=1}^{n} d'_\nu = \mathrm{sgn}(\pi) \cdot \prod_{\nu=1}^{n} d_\nu = \det(A)\,.$$

Nachweis von (6). Es genügt, die Behauptung im Fall benachbarter Zeilen zu beweisen. Sei

$$A = LDPU\,, \quad L = (l_{\nu,\mu})\,, \quad D = \mathrm{diag}(d_1, \ldots, d_n)\,, \quad P = P(\pi)$$

die strikte Bruhat-Zerlegung von A. Wir berechnen eine Bruhat-Zerlegung

$$E_{i \leftrightarrow i+1} \circ A = L'D'P'U'\,, \quad D' = \mathrm{diag}(d'_1, \ldots, d'_n)\,.$$

Zwei Fälle sind zu unterscheiden. Siehe Beispiel 12.6.

(i) *Wir setzen* $l_{i+1,i} = 0$ *oder* $\pi^{-1}(i) > \pi^{-1}(i+1)$ *voraus.* Dann gelten

$$d'_\nu = \begin{cases} d_\nu\,, & \nu \notin \{i, i+1\}\,, \\ d_{i+1}\,, & \nu = i\,, \\ d_i\,, & \nu = i+1\,, \end{cases} \quad , \quad P' = E_{i\leftrightarrow i+1} \circ P\,.$$

Beim Übergang zu $E_{i\leftrightarrow i+1} \circ A$ werden die beiden Diagonalelemente d_i und d_{i+1} vertauscht. Die anderen Diagonalelemente d_ν bleiben erhalten. Die Permutation π wird durch $\pi_{i\leftrightarrow i+1} \circ \pi$ ersetzt. Also folgt

$$\begin{aligned} \det(E_{i\leftrightarrow i+1} \circ A) &= d'_1 \cdot \ldots \cdot d'_n \cdot \operatorname{sgn}(\pi_{i\leftrightarrow i+1} \circ \pi) \\ &= d_1 \cdot \ldots \cdot d_n \cdot (-1) \cdot \operatorname{sgn}(\pi) \\ &= -\det(A)\,. \end{aligned}$$

(ii) *Wir setzen* $l_{i+1,i} \neq 0$ *und* $\pi^{-1}(i) < \pi^{-1}(i+1)$ *voraus.* Dann gelten

$$d'_\nu = \begin{cases} d_\nu\,, & \nu \notin \{i, i+1\}\,, \\ l_{i+1,i}\, d_i\,, & \nu = i\,, \\ -l_{i+1,i}^{-1}\, d_{i+1}\,, & \nu = i+1\,, \end{cases} \quad , \quad P' = P\,.$$

Also gilt $d'_i d'_{i+1} = -d_i d_{i+1}$. Die anderen d_ν bleiben erhalten. Es folgt

$$\begin{aligned} \det(E_{i\leftrightarrow i+1} \circ A) &= d'_1 \cdot \ldots \cdot d'_n \cdot \operatorname{sgn}(\pi) \\ &= (-1) \cdot d_1 \cdot \ldots \cdot d_n \cdot \operatorname{sgn}(\pi) \\ &= -\det(A)\,. \end{aligned}$$

Nachweis von (7). Sei $A = LDPU$ eine Bruhat-Zerlegung von $A \in \mathrm{GL}(n, \mathbb{R})$. Die Matrix B entstehe aus A durch die elementare Zeilenumformung $Z_j + Z_i$. Im Fall $n \geq j > i$ gilt $E_{j;1;i} \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$, denn die i-te Zeile steht oberhalb der j-ten Zeile. Also ist

$$B = E_{j;1;i} A = (E_{j;1;i} L) DPU$$

eine Bruhat-Zerlegung von B. Es folgt

$$\det(B) = \det(A)\,.$$

Der Fall $n \geq i > j$ wird mit Hilfe von (6) auf den schon behandelten Fall zurückgeführt. Es genügt, die i-te und j-te Zeile von A vor und nach der entsprechenden Zeilenoperation zu vertauschen.

Nachweis von (8). Sei $\alpha \in \mathbb{R}^\times$. Die invertierbare Matrix B entstehe aus A durch die elementare Zeilenumformung αZ_i. Sei

$$A = LDPU\,, \quad D = \operatorname{diag}(d_1, \ldots, d_n)$$

eine Bruhat-Zerlegung von A. Dann ist

$$B = E_{\alpha;i} A = (E_{\alpha;i} L E_{\alpha^{-1};i})(E_{\alpha;i} D) PU$$

eine Bruhat-Zerlegung von B. Aus

$$E_{\alpha;i}D = \operatorname{diag}(d_1', \dots, d_n'), \quad d_\nu' = \begin{cases} \alpha d_\nu\,, & \nu = i\,, \\ d_\nu\,, & \nu \neq i \end{cases}$$

folgt

$$\det(B) = \alpha \det(A)\,.$$

Aussage (9) folgt aus (7) und (8).

Nachweis von (10). Wegen (5) gelten die zu (6) bis (9) analogen Aussagen für Spaltenoperationen anstelle von Zeilenoperationen. Wegen (5), (6) und (8) genügt es ferner, die Additivität der Determinante bezüglich der ersten Spalte zu beweisen. Seien $x, y, s_1, \dots, s_n \in \mathbb{R}^n$ gegeben. Es genügt, die Formel

$$\det(x+y, s_2, \dots, s_n) = \det(x, s_2, \dots, s_n) + \det(y, s_2, \dots, s_n)$$

für den Fall zu beweisen, dass die Vektoren $x, s_2, \dots, s_n$ eine Basis des $\mathbb{R}^n$ bilden. In diesem Fall gibt es eindeutig bestimmte Koeffizienten $\alpha, \alpha_2, \dots, \alpha_n \in \mathbb{R}$ mit

$$y = \alpha x + \alpha_2 s_2 + \dots \alpha_n s_n\,.$$

Falls $\alpha = 0$ gilt, folgt die behauptete Formel bereits aus (5) und (9). Wir nehmen daher zusätzlich an, das $\alpha \neq 0$ gilt.

$$\begin{aligned}
\det(x+y, s_2, \dots, s_n) &= \det(x + \alpha x + \alpha_2 s_2 + \dots \alpha_n s_n, s_2, \dots, s_n) \\
&= \det(x + \alpha x, s_2, \dots, s_n) \\
&= (1+\alpha) \det(x, s_2, \dots, s_n) \\
&= \det(x, s_2, \dots, s_n) + \det(\alpha x, s_2, \dots, s_n) \\
&= \det(x, s_2, \dots, s_n) + \det(y, s_2, \dots, s_n)\,.
\end{aligned}$$

Damit ist der Beweis des Satzes beendet. □

Es folgen Beispiele zum Beweis von Aussage (6) des Satzes 12.5 und zur Berechnung der Determinante mit Hilfe von elementaren Umformungen.

Beispiel 12.6. Nach den Ausführungen in 10.2 und 12.4 besteht die strikte Bruhat-Zerlegung $A = LDPU$ mit $P = P(\pi)$ und

$$A = \begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -1 & 4 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Zu Fall (i) *im Beweis von Satz* 12.5*, Aussage* (6)*.* Es gilt

$$\pi^{-1}(1) = 3 > 1 = \pi^{-1}(2).$$

Die Matrix $A' = E_{1\leftrightarrow 2} \circ A$, die aus der Matrix A durch Vertauschung der beiden ersten Zeilen hervorgeht, besitzt die strikte Bruhat-Zerlegung $A' = L'D'P'U'$ mit $L' = (l'_{\nu\mu})$ und

$$A' = \begin{pmatrix} -1 & -3 & -10 \\ 0 & 0 & -2 \\ -4 & -9 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & -21 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & 10 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Offenbar gilt $A = E_{1\leftrightarrow 2} \circ A'$ mit

$$l'_{21} = 0.$$

Der Übergang von A' nach A fällt daher ebenfalls unter den Fall (i).

Zu Fall (ii) *im Beweis von Satz* 12.5*, Aussage* (6)*.* Es gelten

$$\pi^{-1}(2) = 1 < 2 = \pi^{-1}(3), \quad l_{32} = 4 \neq 0.$$

Die Matrix $A'' = E_{2\leftrightarrow 3} \circ A$, die aus A durch Vertauschung der zweiten und ditten Zeile hervorgeht, besitzt die strikte Bruhat-Zerlegung

$$A'' = \begin{pmatrix} 0 & 0 & -2 \\ -4 & -9 & 2 \\ -1 & -3 & -10 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 5 & \frac{1}{4} & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -\frac{3}{4} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & \frac{9}{4} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

□

Beispiel 12.7. Gegeben sei $A \in \mathbb{R}^{4\times 4}$ mit

$$A = \begin{pmatrix} 2 & -2 & 3 & -4 \\ 4 & 2 & 0 & 3 \\ -6 & 2 & -5 & 5 \\ 8 & 0 & 2 & 1 \end{pmatrix}.$$

Wir bringen die Matrix A durch elementare Zeilenumformungen auf obere Dreiecksform. Der Kommentarspalte des zugehörigen Rechenschemas entnehmen wir, welche Änderungen die Determinante bei diesen Umformungen erfahren hat. Diese Änderungen sind bei der Berechnung von $\det(A)$ zu rücksichtigen.

$$\begin{array}{rrrr|l}
2 & -2 & 3 & -4 & \\
4 & 2 & 0 & 3 & \\
-6 & 2 & -5 & 5 & \\
8 & 0 & 2 & 1 & \\
\hline
2 & -2 & 3 & -4 & Z_1 \\
0 & 6 & -6 & 11 & Z_2 - 2Z_1 \\
0 & -4 & 4 & -7 & Z_3 + 3Z_1 \\
0 & 8 & -10 & 17 & Z_4 - 4Z_1 \\
\hline
2 & -2 & 3 & -4 & Z_1 \\
0 & -4 & 4 & -7 & Z_3 \\
0 & 6 & -6 & 11 & Z_2 \\
0 & 8 & -10 & 17 & Z_4 \\
\hline
2 & -2 & 3 & -4 & Z_1 \\
0 & -2 & 2 & -\frac{7}{2} & \frac{1}{2}Z_2 \\
0 & 6 & -6 & 11 & Z_3 \\
0 & 8 & -10 & 17 & Z_4 \\
\hline
2 & -2 & 3 & -4 & Z_1 \\
0 & -2 & 2 & -\frac{7}{2} & Z_2 \\
0 & 0 & 0 & \frac{1}{2} & Z_3 + 3Z_1 \\
0 & 0 & -2 & 3 & Z_4 + 4Z_2 \\
\hline
2 & -2 & 3 & -4 & Z_1 \\
0 & -2 & 2 & -\frac{7}{2} & Z_2 \\
0 & 0 & -2 & 3 & Z_4 \\
0 & 0 & 0 & \frac{1}{2} & Z_3
\end{array}$$

Folglich gilt

$$\det(A) = (-1)^2 \cdot (\tfrac{1}{2})^{-1} \cdot \{\, 2 \cdot (-2) \cdot (-2) \cdot \tfrac{1}{2} \,\} = 8\,.$$

□

Satz 12.8 (Multiplikativität der Determinante)**.** *Sei* $n \in \mathbb{N}$. *Dann gilt*

$$\det(A \circ B) = \det(A) \cdot \det(B)$$

für alle Matrizen A, $B \in \mathbb{R}^{n \times n}$.

Beweis. Die Produktmatrix $A \circ B$ ist genau dann invertierbar, wenn A und B invertierbar sind.

Wenn eine der beiden quadratischen Matrizen A oder B nicht invertierbar ist, dann verschwinden $\det(A \circ B)$ und $\det(A) \cdot \det(B)$.

Seien A, $B \in \mathrm{GL}(n, \mathbb{R})$ gegeben. Wir betrachten zunächst die vier Spezialfälle (1), (2), (3), (4). Der allgemeine Fall (5) wird durch Betrachtung einer Bruhat-Zerlegung der Matrix A behandelt.

(1) Wenn A eine normierte untere Dreiecksmatrix ist, dann gilt

$$\det(A \circ B) = \det(B) = \det(A) \cdot \det(B)\,.$$

Das erste Gleichheitszeichen folgt aus Satz 9.2. Lediglich die normierte untere Dreiecksmatrix einer Bruhat-Zerlegung von B wird beim Übergang zu AB geändert. Die Determinante einer normierten unteren Dreicksmatrix ist gleich 1. Siehe (4) in Satz 12.5. Also gilt auch das zweite Gleichheitszeichen.

(2) Wenn A eine Diagonalmatrix ist, dann gilt

$$\det(A \circ B) = \det(A) \cdot \det(B)\,.$$

Dies folgt aus den Aussagen (4) und (8) des Satzes 12.5. Die Zeilen von B werden mit den entsprechenden Diagonalelementen von A multipliziert.

(3) Wenn A eine Permutationsmatrix ist, dann gilt

$$\det(A \circ B) = \det(A) \cdot \det(B)\,.$$

Beim Übergang von B nach AB werden die Zeilen von B lediglich vertauscht. Siehe Aussage (8) in 9.8. Nach Satz 12.5, Aussage (6) bewirkt die Vertauschung zweier Zeilen einen Vorzeichenwechsel der Determinante.

(4) Wenn A eine normierte obere Dreiecksmatrix ist, dann gilt

$$\det(A \circ B) = \det(B) = \det(A) \cdot \det(B)\,.$$

Das erste Gleichheitszeichen folgt aus Aussage (9) des Satzes 12.5. Dabei verwenden wir den Zerlegungssatz für normierte oberen Dreiecksmatrizen, der aus Satz 9.4 durch Transponieren hervorgeht. Die Determinante einer normierten oberen Dreicksmatrix ist gleich 1. Siehe (4) in Satz 12.5. Also gilt auch das zweite Gleichheitszeichen.

(5) Nun untersuchen wir den allgemeinen Fall. Wenn A eine beliebige invertierbare Matrix ist, betrachten wir eine Bruhat-Zerlegung

$$A = L \circ D \circ P \circ U\,.$$

Dann erhalten wir mit (1), (2), (3) und (4) sukzessive

$$\begin{aligned}
\det(A \circ B) &= \det(L \circ (D \circ P \circ U \circ B)) \\
&= \det(D \circ (P \circ U \circ B)) \\
&= \det(D) \cdot (\det(P \circ U \circ B)) \\
&= \det(D) \cdot \det(P) \cdot \det(U \circ B) \\
&= \det(D) \cdot \det(P) \cdot \det(B) \\
&= \det(A) \cdot \det(B)\,.
\end{aligned}$$

Damit ist der Beweis beendet. □

Satz 12.9 (Laplace. Entwicklung nach der ersten Spalte)**.** *Seien $n \in \mathbb{N}$ und*

$$A = (a_{\nu\mu}) \in \mathbb{R}^{n \times n}$$

beliebig gegeben. Im Fall $n \geq 2$ sei $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ die Matrix, die aus der Matrix A durch Streichen der i-ten Zeile und der j-ten Spalte entsteht. Dann kann die Determinante $\det(A)$ auf folgende Weise induktiv berechnet werden:

$$\det(A) = \begin{cases} a_{11}\,, \quad n = 1\,, \\ a_{11}a_{22} - a_{21}a_{12}\,, \quad n = 2\,, \\ \displaystyle\sum_{\nu=1}^{n} (-1)^{\nu+1} a_{\nu 1} \det(A_{\nu 1})\,, \quad n \geq 2\,. \end{cases}$$

Die Determinante kann auf diese Weise induktiv definiert werden.

Beweis. Für $n = 1$ gilt die Behauptung des Satzes trivialerweise. Für $n = 2$ folgt die Behauptung aus den expliziten Formeln für die strikte Bruhat-Zerlegung in Beispiel 10.7. Sei $n \geq 2$. Nach Satz 12.5 genügt es, die Formel für eine invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ der Form

$$A = \begin{pmatrix} 1 & 0 \\ 0 & A_{11} \end{pmatrix} = \begin{pmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & A_{11} \end{pmatrix}$$

zu beweisen. Eine solche Form kann durch elementare Zeilen- und Spaltenumformungen hergestellt werden. Sei

$$A_{11} = L_{11} \circ D_{11} \circ P_{11} \circ U_{11}$$

eine Bruhat-Zerlegung von A_{11}. Dann ist

$$A = \begin{pmatrix} 1 & 0 \\ 0 & L_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & D_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_{11} \end{pmatrix}$$

eine Bruhat-Zerlegung von A. Offenbar gilt

$$\det(D) = \det(D_{11}) .$$

Sei $\pi \in \mathrm{S}_n$ die Permutation mit $P = P(\pi)$. Dann gilt $\pi(1) = 1$. Sei weiter $\sigma \in \mathrm{S}_{n-1}$ die Permutation mit

$$\sigma(k) = \pi(k+1)$$

für $k = 1, \ldots, n-1$. Dann gilt

$$P_{11} = P(\sigma) .$$

Wegen $\pi(1) = 1$ gilt

$$\frac{\pi(2) - \pi(1)}{2 - 1} \cdot \ldots \cdot \frac{\pi(n) - \pi(1)}{n - 1} = 1 .$$

Wir erhalten

$$\begin{aligned}
\det(P) = \operatorname{sgn}(\pi) &= \prod_{1 \leq \nu < \mu \leq n} \frac{\pi(\mu) - \pi(\nu)}{\mu - \nu} \\
&= \prod_{2 \leq \nu < \mu \leq n} \frac{\pi(\mu) - \pi(\nu)}{\mu - \nu} \\
&= \prod_{1 \leq \nu' < \mu' \leq n-1} \frac{\sigma(\mu') - \sigma(\nu')}{\mu' - \nu'} = \operatorname{sgn}(\sigma) = \det(P_{11})
\end{aligned}$$

Nun folgt

$$\det(A) = \det(A_{11}) .$$

Damit ist der Beweis beendet. □

Satz 12.10 (Entwicklungssatz von Laplace)**.** *Seinen $n \in \mathbb{N}$ mit $n \geq 2$ und*

$$A = (a_{\nu,\mu})_{\nu,\mu=1,\ldots,n} \in \mathbb{R}^{n \times n}$$

beliebig gegeben. Sei $A_{kl} \in \mathbb{R}^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der k-ten Zeile und der l-ten Spalte hervorgeht. Dann gelten die beiden Entwicklungsformeln.

(1) *Entwicklung nach der k-ten Zeile:*

$$\det(A) = \sum_{\mu=1}^{n} (-1)^{k+\mu} a_{k\mu} \det(A_{k\mu}) .$$

(2) *Entwicklung nach der l-ten Spalte:*

$$\det(A) = \sum_{\nu=1}^{n} (-1)^{\nu+l} a_{\nu l} \det(A_{\nu l}) .$$

Beweis. Die Aussagen (1) und (2) folgen aus dem Entwicklungssatz 12.10 und den Eigenschaften 12.5 der Determinante. □

Beispiel 12.11 (Entwicklung nach der zweiten Zeile)**.** Bei Anwendung des Entwicklungssatzes von Laplace ist es günstig, die Determinante nach einer Spalte oder Zeile zu entwickeln, die möglichst viele Nullen enthält.

$$\begin{aligned}\det\begin{pmatrix}-1 & -3 & -10\\ 0 & 0 & -2\\ -4 & -9 & 2\end{pmatrix} &= -0+0-(-2)\cdot\begin{pmatrix}-1 & -3\\ -4 & -9\end{pmatrix}\\ &= 2\cdot(9-12) = -6\,.\end{aligned}$$

Mit Satz 12.5, Aussage (6) folgt

$$\det\begin{pmatrix}0 & 0 & -2\\ -1 & -3 & -10\\ -4 & -9 & 2\end{pmatrix} = (-1)\cdot(-6) = 6\,.$$

Vergleiche Beispiel 12.4. □

Satz 12.12 (Formel von Leibniz)**.** *Seien $n \in \mathbb{N}$ und $A = (a_{\nu\mu})_{\nu,\mu=1,\dots,n} \in \mathbb{R}^{n\times n}$ beliebig gegeben. Dann gilt*

$$\det(A) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\, a_{\pi(1)\,1}\cdots a_{\pi(n)\,n}\,.$$

Die Determinante kann durch die Formel von Leibniz definiert werden.

Beweis. Sei $A = (a_1, \dots, a_n)$. Für $\mu = 1, \dots, n$ gilt

$$a_\mu = \sum_{\nu_\mu=1}^{n} a_{\nu_\mu\mu} e_{\nu_\mu}\,.$$

Der Laufindex ν_μ gehört zur μ-ten Spalte a_μ von A. Die Determinante ist linear in allen Spalten. Ferner gilt $\det(e_{\nu_1}, \dots, e_{\nu_n}) \neq 0$ genau dann, wenn es $\pi \in \mathrm{S}_n$ mit $\pi(\mu) = \nu_\mu$ für alle $\mu = 1, \dots, n$ gibt. Wir erhalten

$$\begin{aligned}\det(A) = \det(a_1,\dots,a_n) &= \sum_{\nu_1,\dots,\nu_n=1}^{n} a_{\nu_1 1}\cdot\ldots\cdot a_{\nu_n n}\cdot\det(e_{\nu_1},\dots,e_{\nu_n})\\ &= \sum_{\pi\in S_n} a_{\pi(1)\,1}\cdot\ldots\cdot a_{\pi(n)\,n}\cdot\det(e_{\pi(1)},\dots,e_{\pi(n)})\\ &= \sum_{\pi\in S_n} a_{\pi(1)\,1}\cdot\ldots\cdot a_{\pi(n)\,n}\cdot\operatorname{sgn}(\pi)\,.\end{aligned}$$

Damit ist der Beweis beendet. □

Beispiel 12.13. Wir betrachten die Spezialfälle $n = 2, 3$.

(1) Sei $n = 2$.

$\pi(1)$	$\pi(2)$	$\mathrm{sgn}(\pi)$
1	2	$+1$
2	1	-1

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \sum_{\pi \in S_2} \mathrm{sgn}(\pi)\, a_{\pi(1)\,1}\, a_{\pi(2)\,2} = a_{11}a_{22} - a_{21}a_{12}\,.$$

(2) Sei $n = 3$.

$\pi(1)$	$\pi(2)$	$\pi(3)$	$\mathrm{sgn}(\pi)$
1	2	3	$+1$
1	3	2	-1
2	1	3	-1
2	3	1	$+1$
3	1	2	$+1$
3	2	1	-1

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{\pi \in S_3} \mathrm{sgn}(\pi)\, a_{\pi(1)\,1}\, a_{\pi(2)\,2}\, a_{\pi(3)\,3}$$

$$\begin{aligned} = \{\, & a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} \\ & - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} \\ & + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}\} \end{aligned}$$

$$= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}.$$

(3) Wir berechnen die Determinante der ersten Matrix aus Beispiel 12.11 mit der Formel von Leibniz.

$$\begin{aligned} &\det \begin{pmatrix} -1 & -3 & -10 \\ 0 & 0 & -2 \\ -4 & -9 & 2 \end{pmatrix} \\ &= \big\{(-1)\cdot 0 \cdot 2 - (-1)(-9)(-2) - 0 \cdot (-3) \cdot 2 \\ &\qquad + 0 \cdot (-9) \cdot (-10) + (-4)(-3)(-2) - (-4) \cdot 0 \cdot (-10)\big\} \\ &= 18 - 24 = -6\,. \end{aligned}$$

□

Beispiel 12.14 (Vektorprodukt)**.** Das *Vektorprodukt* $a \times b$ zweier Vektoren $a = (a_1, a_2, a_3)^t$ und $b = (b_1, b_2, b_3)^t$ des $\mathbb{R}^3$ ist durch

$$a \times b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

definiert. Die rechte Seite fassen wir als *formale* Entwicklung der Determinante

$$\det \begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{pmatrix}$$

nach der dritten Spalte auf, wobei $\{e_1, e_2, e_3\}$ die kanonische Basis des $\mathbb{R}^3$ ist. Im strengen Sinne ist diese Determinante nicht definiert. Wir definieren also zusätzlich

$$\det \begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{pmatrix} = (a_2b_3 - a_3b_2)e_1 - (a_1b_3 - a_3b_1)e_2 + (a_1b_2 - a_2b_1)e_3$$

$$= \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix} .$$

Damit erhalten wir

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \det \begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{pmatrix} .$$

Diese Formel für das Vektorprodukt lässt sich leicht behalten. Solche formale Erweiterungen der Schreibweisen werden aus praktischen und merktechnischen Gründen des Öfteren vorgenommen. Dabei sind allerdings Vorsichtsmaßnahmen bei der Auswertung der Ausdrücke einzuhalten.

Wir können das Vektorprodukt $a \times b$ als Anwendung einer schief-symmetrischen (3×3)-Matrix A auf den Vektor b schreiben. Es gilt

$$\begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix} = a \times b \,.$$

Die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

ist *schief-symmetrisch*, das heißt, es gilt

$$A^t = -A\,.$$

Sämtliche Diagonalelemente einer schief-symmetrischen Matrix verschwinden. Wegen Satz 12.5, Aussage (5) verschwindet auch die Determinante einer schief-symmetrischen Matrix.

Schief-symmetrische (3×3)-Matrizen treten bei der Berechnung der Drehachsen dreidimensionaler Drehungen auf. Siehe Satz und Definition 20.6. □

Beispiel 12.15 (Spatprodukt)**.** Der Vektorraum $\mathbb{R}^3$ sei mit dem euklidischen innneren Produkt $\langle\,.\,,.\,\rangle$ und dem Vektorprodukt $\times$ versehen. Sei $A = (a, b, c) \in \mathbb{R}^3$ die (3×3)-Matrix mit den Spalten $a, b, c \in \mathbb{R}^3$. Dann gilt

$$\det(a, b, c) \equiv \det(A) = \langle a, b \times c\rangle\,.$$

Aus den Eigenschaften der Determinante ergibt sich insbesondere, dass das *Spatprodukt* $\langle a, b \times c\rangle$ unter zyklischen Vertauschungen invariant bleibt:

$$\langle a, b \times c\rangle = \langle b, c \times a\rangle = \langle c, a \times b\rangle\,.$$

Der Absolutbetrag des Spatproduktes der Spaltenvektoren $a, b, c, \in \mathbb{R}^3$ ist das elementargeometrische Volumen $\mathrm{Vol}(P(a, b, c))$ des Parallelflachs

$$P(a, b, c) = \{\,\alpha a + \beta b + \gamma c \mid \alpha, \beta, \gamma \in [0, 1]\}\,,$$

mit den acht Eckpunkten

$$0\,,\ a\,,\ b\,,\ c\,,\ a+b\,,\ a+c\,,\ b+c\,,\ a+b+c\,.$$

Es gilt also

$$\mathrm{Vol}(P(a, b, c)) = |\langle a, b \times c\rangle| = |\det(a, b, c)|\,.$$

□

13 Adjunkte und Regel von Cramer

Satz 13.1. *Sei $n \in \mathbb{N}$. Sei $A = (a_1, \dots, a_n) \in \mathbb{R}^{n\times n}$ eine quadratische Matrix mit den Spalten $a_1, \dots, a_n \in \mathbb{R}^n$. Sei $\varphi_A \colon \mathbb{R}^n \to \mathbb{R}^n$ die lineare Abbildung mit $\varphi_A(x) = Ax$. Dann sind die folgenden Aussagen (1) bis (8) äquivalent.*

(1) *$A = (a_1, \dots, a_n)$ ist invertierbar.*

(2) $\ker(A) = \{0\}$.

(3) $\operatorname{def}(A) = 0$.

(4) $\operatorname{im}(A) = \mathbb{R}^n$.

(5) $\operatorname{rg}(A) = n$.

(6) *$\{a_1, \dots, a_n\}$ ist eine Basis des $\mathbb{R}^n$.*

(7) $\det(A) \neq 0$.

(8) *$\varphi_A \colon \mathbb{R}^n \to \mathbb{R}^n$ ist bijektiv.*

Wenn die äquivalenten Bedingungen (1) bis (8) erfüllt sind, dann gilt

$$\det(A^{-1}) = (\det(A))^{-1} = \frac{1}{\det(A)}\,.$$

Wir ziehen nun eine wichtige Konsequenz aus dem Entwicklungssatz von Laplace. Wenn eine Matrix invertierbar ist, können die Einträge ihrer Inversen als Quotienten geeigneter Determinanten dargestellt werden. Zur Vorbereitung der *Inversionsformel* definieren wir die *Adjunkte* einer quadratischen Matrix. Die *Regel von Cramer* ist eine Konsequenz der Inversionsformel.

Satz und Definition 13.2 (Adjunkte)**.** *Seien $n \in \mathbb{N}$ mit $n \geq 2$ und $A \in \mathbb{R}^{n\times n}$ gegeben.*

(1) *Sei $A_{\nu\mu} \in \mathbb{R}^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der ν-ten Zeile und der μ-ten Spalte hervorgeht. Die Determinante $\det(A_{\nu\mu})$ heißt der (ν,μ)-te* Minor *von A.*

(2) *Die Matrix $\operatorname{adj}(A) = (\alpha_{\nu\mu}) \in \mathbb{R}^{n\times n}$ mit*

$$\alpha_{\nu\mu} = (-1)^{\nu+\mu}\det(A_{\mu\nu})$$

für alle $\nu, \mu = 1, \dots, n$ heißt die Adjunkte *der Matrix A. Hierbei ist auf die Vertauschung der Indizes auf der rechten Seite zu achten.*

(3) *Die Bildung der Adjunkten ist mit der Transposition vertauschbar. Es gilt*

$$(\operatorname{adj}(A))^t = \operatorname{adj}(A^t)\,.$$

(4) *Aus dem Entwicklungssatz von Laplace folgt*

$$A \circ \operatorname{adj}(A) = \operatorname{adj}(A) \circ A = \det(A)\, E_n\,.$$

(5) *Wenn* $\det(A) \neq 0$ *gilt, dann ist* A *invertierbar mit*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)\,.$$

Diese Inversionsformel *folgt aus* (4).

Die Adjunkte kann mit dem Verfahren von Leverrier-Faddeev berechnet werden. Siehe Satz 14.7.

Beweis. Aussage (3) folgt aus der Definition der Adjungierten und den Eigenschaften der Determinante. Wir beweisen (4). Dazu beweisen wir zuerst die Formel

$$\det(A) \cdot \delta_{\nu\mu} = \sum_{l=1}^{n} (-1)^{l+\mu} a_{\nu l} \det(A_{\mu l})\,. \tag{13.1}$$

Im Fall $\nu = \mu$ steht rechts die Entwicklung

$$\sum_{l=1}^{n} (-1)^{l+\mu} a_{\mu l} \det(A_{\mu l})$$

von $\det(A)$ nach der μ-ten Zeile. Dabei wird über den Spaltenindex l summiert. Im Fall $\nu \neq \mu$ ersetzen wir in A die μ-te Zeile durch die ν-te Zeile. In der neuen Matrix A' tritt daher die ν-Zeile zweimal auf. Also gilt $\det(A') = 0$. Die rechte Seite von (13.1) ist die Entwicklung von $\det(A')$ nach der μ-ten Zeile. Damit ist die Formel (13.1) bewiesen. Es folgt

$$\sum_{l=1}^{n} a_{\nu l} \cdot \alpha_{l\mu} = \sum_{l=1}^{n} a_{\nu l} \cdot (-1)^{l+\mu} \det(A_{\mu l}) = \det(A) \cdot \delta_{\nu\mu}\,.$$

Damit ist die erste Gleichung

$$A \circ \operatorname{adj}(A) = \det(A) E_n$$

der Matrixidentität (4) komponentenweise bewiesen. Die zweite Gleichung folgt aus

$$A^t \circ \operatorname{adj}(A^t) = \det(A^t) E_n = \det(A) E_n$$

durch Transponieren. Dabei verwenden wir (3). Die Inversionsformel (5) folgt aus (4). $\square$

Beispiel 13.3. Wir betrachten $A = (a_{\nu\mu}) \in \mathbb{R}^{2\times 2}$. Nach 13.2 gilt

$$\operatorname{adj}(A) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Wenn $\det(A) \neq 0$ gilt, ist A invertierbar mit

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A) = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Siehe Satz 1.1. □

Beispiel 13.4. Wir berechnen die Adjunkte $\operatorname{adj}(A)$ der Matrix

$$A = \begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix}.$$

(1) Minoren berechnen und eintragen.

$$\big(\det(A_{\nu\mu})\big)_{\nu,\mu=1,2,3} = \begin{pmatrix} -96 & -42 & -3 \\ -18 & -8 & 0 \\ -6 & -2 & 0 \end{pmatrix}.$$

(2) Transponieren.

$$\begin{pmatrix} -96 & -18 & -6 \\ -42 & -8 & -2 \\ -3 & 0 & 0 \end{pmatrix}.$$

(3) Vorzeichenschema.

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

(4) Vorzeichen eintragen.

$$\operatorname{adj}(A) = \begin{pmatrix} -96 & 18 & -6 \\ 42 & -8 & 2 \\ -3 & 0 & 0 \end{pmatrix}.$$

(5) Probe mit Hilfe von (4) aus 13.2.

$$\begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix} \begin{pmatrix} -96 & 18 & -6 \\ 42 & -8 & 2 \\ -3 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Insbesondere folgt $\det(A) = 6$. Entwickeln nach der ersten Zeile ergibt

$$\det(A) = 0 + 0 + (-2) \cdot (9 - 12) = 6\,.$$

Siehe Beispiel 14.9. □

Satz 13.5 (Regel von Cramer. Leibniz, Cramer)**.** *Seien $n \in \mathbb{N}$ mit $n \geq 2$ und $A \in \mathbb{R}^{n \times n}$ mit $\det(A) \neq 0$ und $b = (b_\nu) \in \mathbb{R}^n$ gegeben. Dann ist $Ax = b$ ist durch*

$$\xi = (\xi_\nu) = A^{-1}b$$

eindeutig lösbar.

(1) *Aus der Inversionsformel* (4) *des Satzes* 13.2 *folgt*

$$\xi_\nu = \sum_{\mu=1}^{n} (-1)^{\nu+\mu} \frac{\det(A_{\mu\nu})}{\det(A)}\, b_\mu$$

für alle $\nu = 1, \ldots, n$.

(2) *Nach* (1) *gilt*

$$\xi_\nu = \frac{\det(A_\nu)}{\det(A)}$$

für alle $\nu = 1, \ldots, n$. Dabei entsteht $A_\nu \in \mathbb{R}^{n \times n}$ aus A, indem die ν-te Spalte durch b ersetzt wird.

Die Formeln in (1) *und* (2) *zur Berechnung der eindeutig bestimmten Lösung von $Ax = b$ werden* Regel von Cramer *genannt. Formel* (2) *gilt im Fall $n = 1$ trivialerweise. Den Fall $n = 2$ haben wir bereits in Satz* 1.1 *behandelt.*

Beispiel 13.6. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $b \in \mathbb{R}^3$ mit

$$A = (a_1, a_2, a_3) = \begin{pmatrix} -2 & 3 & 1 \\ 6 & -4 & 1 \\ -8 & 37 & 27 \end{pmatrix}, \quad b = \begin{pmatrix} -5 \\ 13 \\ -39 \end{pmatrix}.$$

Die Matrix A ist invertierbar, denn es gilt

$$\det(A) = (+1) \cdot 190 + (-1) \cdot (-50) + (+27) \cdot (-10) = -30\,.$$

Der Einfachheit halber haben wir dabei nach der dritten Spalte entwickelt. Nun berechnen wir den eindeutig bestimmten Vektor $x = (x_1, x_2, x_3)^t \in \mathbb{R}^3$ mit $Ax = b$ mit Hilfe der Regel von Cramer.

$$x_1 = \frac{\det(b, a_2, a_3)}{\det(a_1, a_2, a_3)} = \frac{\det\begin{pmatrix} -5 & 3 & 1 \\ 13 & -4 & 1 \\ -39 & 37 & 27 \end{pmatrix}}{\det\begin{pmatrix} -2 & 3 & 1 \\ 6 & -4 & 1 \\ -8 & 37 & 27 \end{pmatrix}} = \frac{(-120)}{(-30)} = 4\,,$$

$$x_2 = \frac{\det(a_1, b, a_3)}{\det(a_1, a_2, a_3)} = \frac{\det\begin{pmatrix} -2 & -5 & 1 \\ 6 & 13 & 1 \\ -8 & -39 & 27 \end{pmatrix}}{\det\begin{pmatrix} -2 & 3 & 1 \\ 6 & -4 & 1 \\ -8 & 37 & 27 \end{pmatrix}} = \frac{(-60)}{(-30)} = 2\,,$$

$$x_3 = \frac{\det(a_1, a_2, b)}{\det(a_1, a_2, a_3)} = \frac{\det\begin{pmatrix} -2 & 3 & -5 \\ 6 & -4 & 13 \\ -8 & 37 & -39 \end{pmatrix}}{\det\begin{pmatrix} -2 & 3 & 1 \\ 6 & -4 & 1 \\ -8 & 37 & 27 \end{pmatrix}} = \frac{(+90)}{(-30)} = -3\,.$$

□

14 Cayley-Hamilton. Leverrier-Faddeev

Satz und Definition 14.1. *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$ gegeben.*

(1) *$\lambda \in \mathbb{R}$ ist genau dann ein Eigenwert von A, wenn $\det(A - \lambda E_n) = 0$ gilt.*

(2) *Die Menge $\sigma(A)$ der Eigenwerte von A heißt das* Spektrum *der Matrix A.*

(3) *Das Polynom $\chi_A(t) \in \mathbb{R}[t]$ mit*

$$\chi_A(\lambda) = \det(A - \lambda E_n)$$

für alle Ersetzungen der Unbekannten t durch eine beliebige Zahl $\lambda \in \mathbb{R}$ heißt das charakteristische Polynom *von A.*

(4) *Sei $\lambda \in \sigma(A)$. Dann heißt der Teilvektorraum $\ker(A - \lambda E_n) \subseteq \mathbb{R}^n$ der* Eigenraum *zum Eigenwert λ. Die Dimension von $\ker(A - \lambda E_n)$ heißt die* geometrische Vielfachheit *des Eigenwertes λ.*

(5) *Für alle $\lambda \in \mathbb{R}$ gilt $\lambda \in \sigma(A)$ genau dann, wenn λ eine Nullstelle von $\chi_A(t)$ ist, das heißt, wenn $\chi_A(\lambda) = 0$ gilt.*

(6) *Wenn $\lambda \in \mathbb{R}$ eine k-fache Nullstelle von $\chi_A(t)$ ist, wobei $k \in \mathbb{N}$ gilt, dann sagen wir, dass k die* algebraische Vielfachheit *des Eigenwertes λ ist.*

Beispiel 14.2. Wir betrachten die Matrix $A \in \mathbb{R}^{n\times n}$ mit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix}$$

aus Beispiel 8.4. Entwickeln der Matrix

$$A - \lambda E_3 = \begin{pmatrix} 2-\lambda & 0 & 0 \\ 1 & 9-\lambda & -4 \\ 2 & 14 & -6-\lambda \end{pmatrix}$$

nach der ersten Spalte ergibt

$$\chi_A(\lambda) = \det(A - \lambda E_3) = -\lambda^3 + 5\lambda^2 - 8\lambda + 4 = (-1)(\lambda-2)^2(\lambda-1)\,.$$

In Übereinstimmung mit den früheren Rechnungen in 8.4 ist $\lambda_1 = \lambda_2 = 2$ eine doppelte und $\lambda_3 = 1$ eine einfache reelle Nullstelle von $\chi_A(X)$. Im vorliegenden Fall gilt

$$\sigma(A) = \{1, 2\}\,.$$

Außerdem stimmen für jedes $\lambda \in \sigma(A)$ die algebraische und geometrische Vielfachheit von λ überein. Im Allgemeinen gilt dies jedoch nicht. □

Wir erörtern jetzt ein Beispiel von grundlegender Bedeutung. Dieses Beispiel bereitet den Satz von Cayley-Hamilton und das Matrixmodell der komplexen Zahlen vor.

Beispiel 14.3. Es gibt allerdings quadratische reelle Matrizen, die keinen reellen Eigenwert besitzen. Die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

besitzt das charakteristische Polynom

$$\chi_A(\lambda) = \lambda^2 + 1\,.$$

Dieses Polynom besitzt keine reellen Nullstellen. Daher besitzt A keine reellen Eigenwerte. Wenn die Unbekannte t nicht durch reelle Zahlen λ sondern durch die Matrix A selber ersetzt wird, erhalten wir

$$\chi_A(A) = A \circ A + E_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \circ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Nach dem Satz von Cayley-Hamilton gilt, dass jede quadratische reelle Matrix Nullstelle ihres eigenen charakteristischen Polynoms ist. □

Satz 14.4 (Cayley-Hamilton)**.** *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$ gegeben. Dann gilt*

$$\chi_A(A) = 0_{n\times n}\,.$$

Dabei wird die Unbekannte t im charakteristischen Polynom $\chi_A(t) \in \mathbb{R}[t]$ durch die Matrix A ersetzt.

Beweis. Die Aussage des Satzes gilt im Fall $n = 1$ trivialerweise. Wir setzen daher $n \geq 2$ voraus. Es gibt $a_0, \ldots, a_n \in \mathbb{R}$ und $C_0, \ldots, C_{n-1} \in \mathbb{R}^{n\times n}$ mit

$$\chi_A(\lambda) = \sum_{k=0}^{n} a_k \lambda^k\,, \quad \operatorname{adj}(A - \lambda E_n) = \sum_{k=0}^{n-1} \lambda^k C_k$$

für alle $\lambda \in \mathbb{R}$. Nach Satz 13.2 gilt

$$\begin{aligned}\chi_A(\lambda) E_n &= \det(A - \lambda E_n) E_n \\ &= (A - \lambda E_n) \circ \operatorname{adj}(A - \lambda E_n) \\ &= (A - \lambda E_n) \circ (C_0 + \lambda C_1 + \ldots + \lambda^{n-1} C_{n-1})\,.\end{aligned}$$

Ein Koeffizientenvergleich ergibt

$$a_0 E_n = A C_0\,, \quad a_k E_n = A C_k - C_{k-1}\,, \quad a_n E_n = -C_{n-1}$$

für $k = 1, \ldots, n-1$. Eine Multiplikation mit $A^0 = E_n$, A^k respektive A^n liefert

$$a_0 E_n = AC_0\,, \quad a_k A^k = A^{k+1} \circ C_k - A^k \circ C_{k-1}\,, \quad a_n A^n = -A^n \circ C_{n-1}$$

für $k = 1, \ldots, n-1$. Schließlich summieren wir und erhalten

$$\begin{aligned}
\chi_A(A) = \sum_{k=0}^{n} a_k A^k &= AC_0 + (A^2 C_1 - A^1 C_0) \\
&\quad + (A^3 C_2 - A^2 C_1) \\
&\quad \vdots \\
&\quad + (A^n C_{n-1} - A^{n-1} C_{n-2}) - A^n C_{n-1} \\
&= 0\,.
\end{aligned}$$

Damit ist der Beweis beendet. □

Satz und Definition 14.5. *Sei $n \in \mathbb{N}$. Für $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times n}$ wird die* Spur $\mathrm{tr}(A)$ *durch*

$$\mathrm{tr}(A) = \sum_{\nu=1}^{n} a_{\nu\nu} = a_{11} + \ldots + a_{nn}$$

definiert. Seien $\alpha, \beta \in \mathbb{R}$ und $B = (b_{\nu\mu}) \in \mathbb{R}^{n\times n}$. Dann gelten:

(1) $\mathrm{tr}(E_n) = n$.

(2) $\mathrm{tr}(\alpha A + \beta B) = \alpha \cdot \mathrm{tr}(A) + \beta \cdot \mathrm{tr}(B)$.

(3) $\mathrm{tr}(A \circ B) = \mathrm{tr}(B \circ A)$.

Beweis. Die Aussagen (1) und (2) sind klar. Es gilt

$$\mathrm{tr}(AB) = \sum_{\nu=1}^{n} \left(\sum_{\mu=1}^{n} a_{\nu\mu} b_{\mu\nu} \right) = \sum_{\mu=1}^{n} \left(\sum_{\nu=1}^{n} b_{\mu\nu} a_{\nu\mu} \right) = \mathrm{tr}(BA)\,.$$

Damit ist (3) bewiesen. □

Spur und Determinante sind Abbildungen von $\mathbb{R}^{n\times n}$ in $\mathbb{R}$. Sie sind unter allen inneren Automorphismen

$$A \mapsto V^{-1} \circ A \circ V$$

mit $V \in \mathrm{GL}(n, \mathbb{R})$ invariant. Diese Invarianzeigenschaft kann zur Berechnung von $\mathrm{tr}(A)$ und $\det(A)$ ausgenutzt werden. Außerdem bedeutet diese Invarianzeigenschaft, dass Spur und Determinante für lineare Abbildungen $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ unabhängig von einer speziellen Matrixdarstellung einen Sinn machen.

Satz 14.6. *Sei $n \in \mathbb{N}$. Für alle $A \in \mathbb{R}^{n\times n}$ und alle $V \in \mathrm{GL}(n,\mathbb{R})$ gelten die Formeln* (1) *und* (2)*.*

(1) $\mathrm{tr}(V^{-1}AV) = \mathrm{tr}(A)$.

(2) $\det(V^{-1}AV) = \det(A)$.

Die Koeffizienten des charakteristischen Polynoms $\chi_A(t) \in \mathbb{R}[t]$ einer Matrix $A \in \mathbb{R}^{n\times n}$ können sukzessive mit dem Verfahren von Leverrier-Faddeev berechnet werden. Das Verfahren liefert außerdem die Adjunkte $\mathrm{adj}(A)$ und die Determinante $\det(A)$ von A.

Satz 14.7 (Leverrier-Faddeev)**.** *Seien $n \in \mathbb{N}$ mit $n \geq 2$ und $A \in \mathbb{R}^{n\times n}$ gegeben.*

(1) *Die Koeffizienten $p_1, \ldots, p_n \in \mathbb{R}$ des charakteristischen Polynoms*

$$\chi_A(t) = (-1)^n(t^n - p_1t^{n-1} - p_2t^{n-2} - \ldots - p_n) \in \mathbb{R}[t]$$

von A lassen sich mit dem Verfahren von Leverrier-Faddeev *induktiv berechnen.*

(1.1) *Wir setzen $B_0 = E_n$, wobei $E_n \in \mathbb{R}^{n\times n}$ die Einheitsmatrix ist.*

(1.2) *Wir berechnen*

$$A_k = AB_{k-1}\,, \quad p_k = \frac{\mathrm{tr}(A_k)}{k}\,, \quad B_k = A_k - p_kE_n$$

sukzessive für $k = 1, \ldots, n$.

(2) *Nach dem Satz von Cayley-Hamilton gilt $B_n = 0$.*

(3) *Am charakteristischen Polynom $\chi_A(t)$ lassen sich die Spur* $\mathrm{tr}(A)$ *und die Determinante* $\det(A)$ *von A direkt ablesen. Es gelten*

$$\mathrm{tr}(A) = p_1\,, \quad \det(A) = (-1)^{n-1}p_n\,.$$

(4) *Es gilt*

$$\mathrm{adj}(A) = (-1)^{n-1}B_{n-1}\,.$$

(5) *Wenn A invertierbar ist, dann gilt*

$$A^{-1} = \frac{1}{p_n}\,B_{n-1}\,.$$

Eigenschaft (2) *dient zur Kontrolle der Rechnung.*

Beispiel 14.8 (Fortsetzung von Beispiel 14.2)**.** Wir untersuchen die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix}$$

mit dem Verfahren von Leverrier-Faddeev. Die Rechnung fassen wir in folgendem Schema zusammen.

$$A_1 = AE_3 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix}, \quad p_1 = 5, \quad B_1 = \begin{pmatrix} -3 & 0 & 0 \\ 1 & 4 & -4 \\ 2 & 14 & -11 \end{pmatrix}.$$

$$A_2 = \begin{pmatrix} -6 & 0 & 0 \\ -2 & -20 & 8 \\ -4 & -28 & 10 \end{pmatrix}, \quad p_2 = -8, \quad B_2 = \begin{pmatrix} 2 & 0 & 0 \\ -2 & -12 & 8 \\ -4 & -28 & 18 \end{pmatrix}.$$

$$A_3 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad p_3 = 4, \quad B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\chi_A(t) = (-1)^3(t^3 - 5t^2 + 8t - 4) = (-1)(t-2)^2(t-1).$$

$$\det(A) = (-1)^2\, p_3 = 4.$$

$$A^{-1} = \frac{1}{p_3} B_2 = \frac{1}{4} \begin{pmatrix} 2 & 0 & 0 \\ -2 & -12 & 8 \\ -4 & -28 & 18 \end{pmatrix}.$$

□

Beispiel 14.9 (Fortsetzung von Beispiel 13.4)**.** Wir berechnen die Adjunkte adj(A) und die Determinante $\det(A)$ der Matrix

$$A = \begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix}$$

mit dem Verfahren von Leverrier-Faddeev. Wie in 13.4 erhalten wir

$$\operatorname{adj}(A) = \begin{pmatrix} -96 & 18 & -6 \\ 42 & -8 & 2 \\ -3 & 0 & 0 \end{pmatrix}, \quad \det(A) = 6.$$

Die Rechnung fassen wir in folgendem Schema zusammen.

$$A_1 = A = \begin{pmatrix} 0 & 0 & -2 \\ -1 & -3 & -10 \\ -4 & -9 & 2 \end{pmatrix},$$

$$p_1 = \mathrm{tr}(A_1) = -1\,, \quad B_1 = A_1 - p_1 E_3 = \begin{pmatrix} 1 & 0 & 2 \\ -1 & -2 & -10 \\ 4 & -9 & 3 \end{pmatrix}.$$

$$A_2 = AB_1 = \begin{pmatrix} 8 & 18 & -6 \\ 42 & 96 & 2 \\ -3 & 0 & 104 \end{pmatrix},$$

$$p_2 = \frac{\mathrm{tr}(A_2)}{2} = 104\,, \quad B_2 = A_2 - p_2 E_3 = \begin{pmatrix} -96 & 18 & -6 \\ 42 & -8 & 2 \\ -3 & 0 & 0 \end{pmatrix}.$$

$$A_3 = AB_2 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}, \quad p_3 = \frac{\mathrm{tr}(A_3)}{3} = 6\,, \quad B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\begin{aligned} \chi_A(t) &= (-1)^n (t^n - p_1 t^{n-1} - p_2 t^{n-2} - \ldots - p_n) \\ &= (-1)^3 (t^3 + t^2 - 104t - 6) \\ &= -t^3 - t^2 + 104t + 6\,. \end{aligned}$$

$$\mathrm{adj}(A) = (-1)^2 B_2 = \begin{pmatrix} -96 & 18 & -6 \\ 42 & -8 & 2 \\ -3 & 0 & 0 \end{pmatrix}.$$

$$\det(A) = (-1)^2\, p_3 = 6\,.$$

□

Beispiel 14.10. Wir berechnen das charakteristische Polynom $\chi_A(t)$ explizit in den Fällen $n = 2$ und $n = 3$.

(1) Für alle $A \in \mathbb{R}^{2\times 2}$ gilt

$$\chi_A(t) = \det(A - tE_2) = (-1)^2 (t^2 - \mathrm{tr}(A) \cdot t + \det(A))\,.$$

(2) Für alle $A \in \mathbb{R}^{3\times 3}$ gilt

$$\chi_A(t) = \det(A - tE_3) = (-1)^3(t^3 - \operatorname{tr}(A) \cdot t^2 + \operatorname{tr}(\operatorname{adj}(A)) \cdot t - \det(A))\,.$$

Im Fall $n = 3$ tritt die Spur der Adjunkten in der Formel für das charakteristische Polynom auf. Wie das Beispiel

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix}, \quad \operatorname{tr}(A) = 5\,, \quad \det(A) = 4\,,$$

$$\operatorname{adj}(A) = \begin{pmatrix} 2 & 0 & 0 \\ -2 & -12 & 8 \\ -4 & -28 & 18 \end{pmatrix}, \quad \operatorname{tr}(\operatorname{adj}(A)) = 8$$

zeigt, stimmen die Spur von A und die Spur von $\operatorname{adj}(A)$ im Allgemeinen nicht überein. Siehe 14.8. Dagegen gilt für $A \in \mathbb{R}^{2\times 2}$ stets $\operatorname{tr}(A) = \operatorname{tr}(\operatorname{adj}(A))$. Siehe 13.3. Wenn wir in (1) und (2) für die Unbekannte t die Matrix A einsetzen, erhalten wir mit dem Satz von Cayley-Hamilton die folgenden Aussagen (3) und (4).

(3) Für alle $A \in \mathbb{R}^{2\times 2}$ gilt

$$A^2 = \operatorname{tr}(A) \cdot A - \det(A) \cdot E_2\,.$$

(4) Für alle $A \in \mathbb{R}^{3\times 3}$ gilt

$$A^3 = \operatorname{tr}(A) \cdot A^2 - \operatorname{tr}(\operatorname{adj}(A)) \cdot A + \det(A) \cdot E_3\,.$$

Wir betrachten noch einmal die Matrix A aus 14.8. Ausmultiplizieren ergibt

$$A^2 = \begin{pmatrix} 4 & 0 & 0 \\ 3 & 25 & -12 \\ 6 & 42 & -20 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 8 & 0 & 0 \\ 7 & 57 & -28 \\ 14 & 98 & -48 \end{pmatrix}.$$

Offenbar gilt

$$\begin{pmatrix} 8 & 0 & 0 \\ 7 & 57 & -28 \\ 14 & 98 & -48 \end{pmatrix} = 5 \begin{pmatrix} 4 & 0 & 0 \\ 3 & 25 & -12 \\ 6 & 42 & -20 \end{pmatrix} - 8 \begin{pmatrix} 2 & 0 & 0 \\ 1 & 9 & -4 \\ 2 & 14 & -6 \end{pmatrix} + 4 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

□

Nach dem folgenden Satz ist jedes reelle Polynom vom Grad $n \in \mathbb{N}$, dessen Leitkoeffizienten gleich $(-1)^n$ ist, das charakteristische Polynom einer Matrix aus $\mathbb{R}^{n\times n}$.

Satz 14.11 (Begleitmatrix)**.** *Seien $n \geq N$ mit $n \geq 2$ und $p_1, \ldots, p_n \in \mathbb{R}$ beliebig gegeben. Dann ist das Polynom $\varphi(t) \in \mathbb{R}[t]$ mit*

$$\varphi(t) = (-1)^n (t^n - p_1 t^{n-1} - p_2 t^{n-2} - \ldots - p_n)$$

das charakteristische Polynom der Matrix $A \in \mathbb{R}^{n \times n}$ mit

$$A = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & 0 & p_n \\ 1 & 0 & 0 & \ldots & 0 & 0 & 0 & p_{n-1} \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 & p_2 \\ 0 & 0 & 0 & \ldots & 0 & 0 & 1 & p_1 \end{pmatrix}.$$

Diese Schreibweise ist so zu verstehen, dass in der ersten Diagonale unterhalb der Hauptdiagonalen der Matrix A alle Einträge gleich 1 sind, in der letzten Spalte von oben nach unten die Koeffizienten $p_n, \ldots, p_1$ eingetragen werden, und alle anderen Einträge gleich 0 sind. Die Matrix A heißt die Begleitmatrix *des Polynoms $\varphi(t)$.*

Beispiel 14.12 (Fortsetzung des fundamentalen Beispiels 14.3)**.** Die Begleitmatrix des Polynoms

$$\varphi(t) = t^2 + 1 = (-1)^2(t^2 - p_1 t - p_2) \in \mathbb{R}[t]$$

mit $p_1 = 0$ und $p_2 = -1$ ist die Matrix

$$A = \begin{pmatrix} 0 & p_2 \\ 1 & p_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Wir verweisen auf Satz 15.5. Im Matrixmodell der komplexen Zahlen ist diese Matrix die imaginäre Einheit i. Die Matrizen $\pm A$ sind die einzigen Matrizen in $\mathbb{R}^{2 \times 2}$, die die Matrixgleichung

$$A^2 + E_2 = 0_{2 \times 2}$$

erfüllen. Die Idee ist, die fehlenden reellen Wurzeln der Gleichung $t^2 + 1 = 0$ als reelle (2×2)-Matrizen zu konstruieren. □

Beispiel 14.13.

(1) Für alle $A \in \mathbb{R}^{2 \times 2}$ gilt

$$\det(E_2 + tA) = 1 + \operatorname{tr}(A) \cdot t + \det(A) \cdot t^2 .$$

(2) Für alle $A \in \mathbb{R}^{3 \times 3}$ gilt

$$\det(E_3 + tA) = 1 + \operatorname{tr}(A) \cdot t + \operatorname{tr}(\operatorname{adj}(A)) \cdot t^2 + \det(A) \cdot t^3 .$$

□

Beispiel 14.14. Sei $n \in \mathbb{N}$ mit $n \geq 2$. Für alle $A \in \mathbb{R}^{n \times n}$ gilt

$$\frac{d}{dt} \det(E_n + tA) = \operatorname{tr}(A \circ \operatorname{adj}(E + tA)) .$$

Dabei wird die Ableitung eines reellen Polynoms rein algebraisch durch

$$\frac{d}{dt}\left(\sum_{\mu=0}^{m} a_\mu t^\mu\right) = \begin{cases} 0 , & \mu = 0 , \\ \displaystyle\sum_{\mu=1}^{m-1} \mu a_\mu t^{\mu-1} , & m \in \mathbb{N} \end{cases}$$

definiert. □

15 Matrixmodell der komplexen Zahlen

Satz und Definition 15.1. *Ein Tripel* $(K, +, \cdot)$, *das aus einer nicht-leeren Menge* K *und zwei Abbildungen* $+ : K \times K \to K$ *mit* $(x, y) \mapsto x + y$ *und* $\cdot : K \times K \to K$ *mit* $(x, y) \mapsto xy = x \cdot y$ *besteht, heißt ein* Körper, *wenn die folgenden Bedingungen* (1) *bis* (4) *erfüllt sind. Dann nennen wir die Abbildung* $+$ *die* Addition *und die Abbildung* $\cdot$ *die* Multiplikation *des Körpers* $(K, +, \cdot)$.

(1) $(\forall x, y \in K: \quad x + y = y + x\,, \quad xy = yx\,.$

(2) $(\forall x, y, z \in K): \quad x + (y + z) = (x + y) + z\,, \quad x(yz) = (xy)z\,.$

(3) $(\forall x, y, z \in K): \quad x(y + z) = xy + xz\,.$

(4) *Es gibt Elemente* $0 \in K$ *und* $1 \in K$ *derart, dass* (4.1) *bis* (4.5) *gelten.*

 (4.1) $0 \neq 1\,.$

 (4.2) $(\forall x \in K): \quad x + 0 = x\,.$

 (4.3) $(\forall x \in K)(\exists y \in K): \quad x + y = 0\,.$

 (4.4) $(\forall x \in K): \quad 1 \cdot x = x\,.$

 (4.5) $(\forall x \in K \setminus \{0\})(\exists y \in K): \quad xy = 1\,.$

Es gelten die folgenden Einzigkeitsaussagen (5), (6), (7).

(5) *Die Elemente* 0 *und* 1 *sind bereits durch die Bedingungen* (1), (2), (4.2), (4.4) *eindeutig bestimmt. Diese Elemente heißen die* Null *respektive die* Eins *des Körpers* $(K, +, \cdot)$.

(6) *Die Elemente* $y \in K$ *mit* $x + y = 0$ *sind eindeutig durch* $x \in K$ *bestimmt. Sie heißen das* negative Element $(-x)$ *zu* x.

(7) *Die Elemente* $y \in K \setminus \{0\}$ *mit* $xy = 1$ *sind durch* $x \in K \setminus \{0\}$ *eindeutig bestimmt. Sie heißen das* inverse *oder* reziproke *Element* x^{-1} *zu* x.

Wir nennen die Menge $K^\times = K \setminus \{0\}$ *die* Menge invertierbaren Elemente *des Körpers* $(K, +, \cdot)$.

Wenn keine Verwechslungen zu befürchten sind, schreiben wir K anstelle von $(K, +, \cdot)$. Außerdem unterscheiden wir in der Notation zwischen den Additionen und Multiplikationen verschiedener Körper nicht.

Beispiele 15.2.

(1) Das Tripel $(\mathbb{Q}, +, \cdot)$ bestehend aus der Menge $\mathbb{Q}$ der rationalen Zahlen und der üblichen Addition und der üblichen Multiplikation ist ein Körper.

(2) Das Tripel $(\mathbb{R}, +, \cdot)$ bestehend aus der Menge $\mathbb{R}$ der reellen Zahlen und der üblichen Addition und der üblichen Multiplikation ist ein Körper.

□

Definition 15.3. *Sei K ein Körper. Ein Paar (L, Φ) bestehend aus einem Körper L und einer Abbildung $\Phi : K \to L$ heißt eine* Körpererweiterung *von K, wenn die folgenden Bedingungen* (1), (2), (3) *erfüllt sind.*

(1) *Für alle $x,\ y \in K$ gilt $\Phi(x + y) = \Phi(x) + \Phi(y)$.*

(2) *Für alle $x,\ y \in K$ gilt $\Phi(x \cdot y) = \Phi(x) \cdot \Phi(y)$.*

(3) *Aus $\Phi(x) = \Phi(y)$ folgt $x = y$ für alle $x,\ y \in K$.*

In dieser Situation sagen wir auch, dass L eine Körpererweiterung von K ist. Manchmal sagen wir, dass Φ eine Körpererweiterung von K ist. Außerdem ist es üblich, K mit der Bildmenge $\Phi(K)$ zu identifizieren. Wir sagen dann, dass K ein Teilkörper *von L ist.*

Satz 15.4. *Sei K eine Körper. Für eine Körpererweiterung (L, Φ) von K gelten die folgenden Aussagen* (1), (2), (3).

(1) $\Phi(0) = 0$.

(2) $\Phi(K^\times) \subseteq L^\times$.

(3) $\Phi(1) = 1$.

Wir konstruieren in Satz 15.5 das *Matrixmodell* des Körpers $\mathbb{C}$ der komplexen Zahlen. Die komplexen Zahlen werden als reelle (2×2)-Matrizen der

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Form konstruiert. Der Körper $\mathbb{C}$ ist eine Körpererweiterung der Körpers $\mathbb{R}$ der reellen Zahlen. Dabei werden die reellen Zahlen a mit den Diagonalmatrizen

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

identifiziert. Wir erinnern an die Beispiele 14.3 und 14.12 und den Satz 14.4 von Cayley-Hamilton. Im Körper $\mathbb{C}$ ist die Gleichung $x^2 + 1 = 0$ lösbar. Nach dem Fundamentalsatz 15.12 besitzt jedes komplexe Polynom n-ten Grades n Nullstellen.

Satz 15.5 (Matrixmodell der komplexen Zahlen)**.** *Sei $\mathbb{C} \subseteq \mathbb{R}^{2\times 2}$ die Menge aller reellen (2×2)-Matrizen von der Form*

$$c = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

mit $a, b \in \mathbb{R}$. Dann gelten die folgenden Aussagen (1) *bis* (8).

(1) *Die Menge $\mathbb{C}$ enthält die Nullmatrix $0_{2\times 2}$ und die Einheitsmatrix E_2.*

(2) *Für alle* a_1, a_2, b_1, b_2 *gilt*

$$\begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} + \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 & -(b_1 + b_2) \\ b_1 + b_2 & (a_1 + a_2) \end{pmatrix} .$$

Also liegen Summen von Matrizen aus $\mathbb{C}$ *ebenfalls in* $\mathbb{C}$.

(3) *Für alle Matrizen* $c \in \mathbb{C}$ *liegen auch die negativen Matrizen* $-c$ *in* $\mathbb{C}$.

(4) *Für alle* a_1, a_2, b_1, b_2 *gilt*

$$\begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} \circ \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix} = \begin{pmatrix} a_1 a_2 - b_1 b_2 & -(a_1 b_2 + a_2 b_1) \\ a_1 b_2 + a_2 b_1 & (a_1 a_2 - b_1 b_2) \end{pmatrix} .$$

Also liegen Produkte von Matrizen aus $\mathbb{C}$ *ebenfalls in* $\mathbb{C}$.

(5) *Eine Matrix*

$$c = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{C}$$

ist genau dann invertierbar, wenn $\det(c) = a^2 + b^2 \neq 0$ *gilt. Dann gilt*

$$c^{-1} = \frac{1}{a^2 + b^2} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in \mathbb{C} .$$

Also liegen für alle Matrizen $c \in \mathbb{C} \setminus \{0\}$ *auch die inversen Matrizen* c^{-1} *in* $\mathbb{C}$.

(6) *Das Tripel* $(\mathbb{C}, +, \circ)$ *ist ein Körper. Dabei ist die Addition* $+$ *die Matrizenaddition und die Multiplikation* $\circ$ *die Matrizenmultiplikation in* $\mathbb{R}^{2\times 2}$.

(6.1) *Das Tripel* $(\mathbb{C}, +, \circ)$ *heißt* Körper der komplexen Zahlen.

(6.2) *Die Elemente von* $\mathbb{C}$ *heißen* komplexe Zahlen.

(6.3) *Das Produkt* $c_1 \circ c_2$ *zweier komplexer Zahlen* c_1 *und* c_2 *schreiben wir wie üblich in der Form* $c_1 c_2 = c_1 \cdot c_2$.

(7) *Die Abbildung* $\Phi : \mathbb{R} \to \mathbb{C}$ *mit*

$$\Phi : a \mapsto \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

ist eine Körpererweiterung.

(8) *Im Körper* $\mathbb{C}$ *ist die Gleichung* $x^2 = -1$ *lösbar. Es gilt*

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Nach dem Fundamentalsatz der Algebra zerfällt jedes komplexe Polynom vollständig in Linearfaktoren. Siehe Satz 15.12.

Satz und Definition 15.6 (Schreibweisen)**.** *Jede komplexe Zahl*

$$c = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{C}$$

lässt sich mit $a, b \in \mathbb{R}$ auf genau eine Weise in der Form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$= a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

zerlegen. Wir vereinfachen die Schreibweise, indem wir eine reelle Zahl mit ihrem Bild unter der Körpererweiterung $\Phi : \mathbb{R} \to \mathbb{C}$ identifizieren.

(1) *Wir setzen*

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad a + bi = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Wie üblich lassen wir den Punkt für die Multiplikation fort. Wegen der Kommutativität der Multiplikation in $\mathbb{C}$ gilt insbesondere $bi = ib$.

(2) *Die komplexe Zahl i ist die* imaginäre Einheit*. Es gilt*

$$(\pm i)^2 = -1\,.$$

(3) *Die reellen Zahlen a und b heißen der* Real- *respektive der* Imaginärteil *der komplexen Zahl $c = a + bi$.*

$$c = a + bi\,, \quad \mathrm{Re}(c) = a\,, \quad \mathrm{Im}(c) = b\,.$$

(4) *Die komplexe Zahl $\bar{c} = a - bi$ heißt* konjugiert komplexe *Zahl zu c.*

$$\bar{c} = a - bi\,, \quad \mathrm{Re}(\bar{c}) = \mathrm{Re}(c) = a\,, \quad \mathrm{Im}(\bar{c}) = -\mathrm{Im}(c) = -b\,.$$

(5) *Für alle $a_1, a_2, b_1, b_2 \in \mathbb{R}$ gilt*

$$(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + i(b_1 + b_2)\,.$$

(6) *Für alle $a_1, a_2, b_1, b_2 \in \mathbb{R}$ gilt*

$$(a_1 + b_1 i)(a_2 + b_2 i) = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1) i\,.$$

(7) *Insbesondere ergibt sich*

$$\bar{c}c = (a - bi)(a + bi) = a^2 + b^2 \geq 0\,.$$

Die nicht-negative Quadratwurzel

$$|c| = \sqrt{\bar{c}c} = \sqrt{a^2 + b^2} \geq 0$$

heißt Absolutbetrag *der komplexen Zahl $c = a + bi$.*

Beispiel 15.7.

$$(1-2i)+(3+4i)=4+2i\,.$$

$$(1-2i)-(3+4i)=(-2)+(-6)\cdot i=-2-6i=(-2)(1+3i)\,.$$

$$(1-2i)\cdot(3+4i)=11-2i\,.$$

$$\overline{(3+4i)}=3-4i\,,\quad |3+4i|^2=\overline{(3+4i)}\cdot(3+4i)=25\,.$$

$$|3\pm 4i|=5\,.$$

$$\frac{1-2i}{3+4i}=\frac{1-2i}{3+4i}\cdot\frac{3-4i}{3-4i}=\frac{-5-10i}{25}=-\frac{1-2i}{5}=-\frac{1}{5}-\frac{2}{5}\,i\,.$$

□

Satz 15.8 (Polardarstellung komplexer Zahlen)**.**

(1) *Jede komplexe Zahl $c\in\mathbb{C}^\times\subseteq\mathrm{GL}(2,\mathbb{R})$ besitzt eine Darstellung der Form*

$$c=|c|\begin{pmatrix}\cos(\varphi) & -\sin(\varphi)\\ \sin(\varphi) & \cos(\varphi)\end{pmatrix}$$

mit einem eindeutig bestimmten reellen Parameter $\varphi\in(-\pi,\pi]$. Diese Darstellung heißt die Polardarstellung *von c.*

(2) *Eine komplexe Zahlen $c\in\mathbb{C}^\times\subseteq\mathrm{GL}(2,\mathbb{R})$ ist demnach eine ebene Drehstreckung mit Streckungsfaktor $|c|>0$ und Drehwinkel $\varphi\in(-\pi,\pi]$. Es gelten*

$$|c|=\sqrt{\det(c)}\,,\quad \cos(\varphi)=\frac{\mathrm{tr}(c)}{2\sqrt{\det(c)}}=\frac{\mathrm{Re}(c)}{|c|}\,,\quad \sin(\varphi)=\frac{\mathrm{Im}(c)}{|c|}\,.$$

Dabei ist $\varphi\in(-\pi,\pi]$ durch die beiden letzten Gleichungen eindeutig bestimmt.

Beispiel 15.9.

$$1-i=\sqrt{2}\left(\frac{1}{\sqrt{2}}-\frac{i}{\sqrt{2}}\right)=\sqrt{2}\left(\cos\left(-\frac{\pi}{4}\right)+i\sin\left(-\frac{\pi}{4}\right)\right)\,.$$

□

Tabelle 15.10.

φ	$\frac{\pi}{6}$	$\frac{\pi}{5}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{4\pi}{5}$	$\frac{5\pi}{6}$
$\cos(\varphi)$	$\frac{\sqrt{3}}{2}$	$\frac{1+\sqrt{5}}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{1+\sqrt{5}}{4}$	$-\frac{\sqrt{3}}{2}$
$\sin(\varphi)$	$\frac{1}{2}$	$\frac{\sqrt{10-2\sqrt{5}}}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{10-2\sqrt{5}}}{4}$	$\frac{1}{2}$

□

Aus der Multiplikationsregel für komplexe Zahlen und den Additionstheoremen für Cosinus und Sinus ergibt sich der folgende Satz 15.11.

Satz 15.11 (Euler-Moivre)**.**

(1) *Für alle $\varphi, \psi \in \mathbb{R}$ gilt*

$$(\cos(\varphi) + i\sin(\varphi))(\cos(\psi) + i\sin(\psi)) = \cos(\varphi + \psi) + i\sin(\varphi + \psi)\,.$$

(2) *Für alle $n \in \mathbb{N}_0$ und alle $\varphi \in \mathbb{R}$ gilt*

$$(\cos(\varphi) + i\sin(\varphi))^n = \cos(n\varphi) + i\sin(n\varphi)\,.$$

Die Aussagen (1) *und* (2) *sind eine elegante Zusammenfassung der Additionstheoreme.*

Nach Satz 15.11 ist eine Exponentialschreibweise sinnvoll. Wir setzen

$$e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)\,, \quad \varphi \in \mathbb{R}\,.$$

Speziell für $\varphi = \pi$ erhalten wir die berühmte *Euler'sche Formel*

$$e^{i\pi} + 1 = 0\,.$$

In der Analysis werden die Exponentialfunktion, der Cosinus und der Sinus durch eine Potenzreihe definiert. Dann wird aus der obigen Definitionsgleichung für $e^{i\varphi}$ eine beweisbare Gleichung. Aus Satz 15.11 folgen die Rechenregeln

$$e^{i\varphi}e^{i\psi} = e^{i(\varphi+\psi)}\,, \quad (e^{i\varphi})^n = e^{in\varphi}\,, \quad |e^{i\varphi}| = 1$$

für alle $\varphi, \psi \in \mathbb{R}$ und alle $n \in \mathbb{N}_0$.

Wir formulieren die Polardarstellung mit der Exponentialschreibweise um. Sei $c \in \mathbb{C}^\times \subseteq \mathrm{GL}(2,\mathbb{R})$. Dann gilt

$$c = |c|\begin{pmatrix}\cos(\varphi) & -\sin(\varphi)\\ \sin(\varphi) & \cos(\varphi)\end{pmatrix} = |c|(\cos(\varphi) + i\sin(\varphi)) = |c|\,e^{i\varphi}$$

mit einem eindeutig bestimmten $\varphi \in (-\pi, \pi]$.

Satz 15.12 (Fundamentalsatz der Algebra. Gauß)**.** *Jedes normierte komplexe Polynom* $p(x) \in \mathbb{C}[x]$ *vom Grade* $n \in \mathbb{N}$ *besitzt eine Zerlegung*

$$p(x) = \prod_{k=1}^{n} (x - \lambda_k)$$

in ein Produkt aus Linearfaktoren $(x - \lambda_k)$ *mit* $\lambda_k \in \mathbb{C}$*. Diese Zerlegung ist bis auf die Reihenfolge der Linearfaktoren eindeutig bestimmt.*

Satz 15.13 (Auflösungsformeln für quadratische Gleichungen)**.**

(1) *Sei* $w \in \mathbb{C}$ *beliebig gegeben. Dann gilt* $|w| \pm \mathrm{Re}(w) \geq 0$*. Die Formeln*

$$\zeta = \begin{cases} \pm \dfrac{\sqrt{|w| + \mathrm{Re}(w)} + i\,\sqrt{|w| - \mathrm{Re}(w)}}{\sqrt{2}}\,, & \mathrm{Im}(w) \geq 0\,, \\ \\ \pm \dfrac{\sqrt{|w| + \mathrm{Re}(w)} - i\,\sqrt{|w| - \mathrm{Re}(w)}}{\sqrt{2}}\,, & \mathrm{Im}(w) < 0 \end{cases}$$

liefern alle Lösungen $\zeta \in \mathbb{C}$ *der quadratischen Gleichung* $z^2 = w$*.*

(2) *Seien* $p,\ q \in \mathbb{C}$ *beliebig gegeben. Sei* $\zeta \in \mathbb{C}$ *mit*

$$\zeta^2 = \left(\frac{p}{2}\right)^2 - q$$

gegeben. Dann sind

$$z_1 = -\frac{p}{2} + \zeta\,, \quad z_2 = -\frac{p}{2} - \zeta$$

die Lösungen der quadratischen Gleichung $z^2 + pz + q = 0$*.*

16 Spektralsatz. Projektionen. Gram-Schmidt

Wir betrachten in diesem Abschnitt symmetrische reelle Matrizen. Nach dem Spektralsatz 16.6 lassen sich diese Matrizen als endliche Linearkombinationen von Projektionsmatrizen darstellen.

Satz und Definition 16.1. *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\, .\, , .\, \rangle$ und der euklidischen Norm $\|\, .\, \|$ versehen.*

(1) $A \in \mathbb{R}^{n\times n}$ *heißt* positiv, *wenn es* $B \in \mathbb{R}^{n\times n}$ *mit* $A = B^t B$ *gibt.*

(2) *Jede positive Matrix* $A \in \mathbb{R}^{n\times n}$ *ist symmetrisch.*

(3) *Eine Matrix* $P \in \mathbb{R}^{n\times n}$ *mit* $P = P^t = P^2$ *heißt eine* Projektionsmatrix.

(4) $\mathrm{Proj}(n, \mathbb{R}) = \{P \in \mathbb{R}^{n\times n} \mid P = P^t = P^2\}$.

(5) *Jede Projektionsmatrix ist positiv.*

Satz und Definition 16.2 (Projektionssatz)**.** *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\, .\, , .\, \rangle$ und der euklidischen Norm $\|\, .\, \|$ versehen. Sei $\mathcal{E}_n = \{e_1, \ldots, e_n\}$ die kanonische Basis des $\mathbb{R}^n$ Sei $X \subseteq \mathbb{R}^n$ ein beliebiger Teilvektorraum.*

(1) *Die Teilmenge* $X^\perp \subseteq \mathbb{R}^n$ *mit*

$$X^\perp = \{\, x \in \mathbb{R}^n \mid (\forall v \in X) : \ \langle x, v\rangle = 0\,\}$$

ist ein Teilvektorraum.

(1.1) $X^\perp$ *heißt der* Orthogonalraum zu X.

(1.2) $X^{\perp\perp} = (X^\perp)^\perp$ *heißt der* Biorthogonalraum zu X.

(1.3) *Jeder Vektor* $x \in \mathbb{R}^n$ *besitzt eine eindeutig bestimmte Zerlegung*

$$x = x' + x''$$

mit $x' \in X$ *und* $x'' \in X^\perp$.

(1.4) $X^{\perp\perp} = X$.

(2) *Die Abbildung* $\pi : \mathbb{R}^n \to \mathbb{R}^n$ *mit*

$$\pi(x) = x'$$

ist linear. Wir nennen π *die* orthogonale Projektion des $\mathbb{R}^n$ auf X. *Sei*

$$P = M_{\mathcal{E}_n\mathcal{E}_n}(\pi) \in \mathbb{R}^{n\times n}$$

die Matrixdarstellung von π *bezüglich* $\mathcal{E}_n$. *Es gelten:*

(2.1) $(\forall x \in \mathbb{R}^{n\times n})$: $Px = \pi(x) = x'$.

(2.2) $P \in \mathrm{Proj}(n, \mathbb{R})$.

(2.3) $X = \mathrm{im}(\pi) = \mathrm{im}(P)$.

(2.4) $X^\perp = \ker(\pi) = \ker(P)$.

(2.5) *P ist die einzige Projektionsmatrix $M \in \mathrm{Proj}(n, \mathbb{R})$ mit $X = \mathrm{im}(M)$.*

(3) *Die Abbildung $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ mit*

$$\sigma(x) = x''$$

ist die orthogonale Projektion des $\mathbb{R}^n$ auf den Orthogonalraum $X^\perp$. Sei

$$Q = M_{\mathcal{E}_n \mathcal{E}_n}(\sigma) \in \mathbb{R}^{n \times n}$$

die Matrixdarstellung von σ bezüglich $\mathcal{E}_n$. Es gelten:

(3.1) $(\forall x \in \mathbb{R}^n)\colon \sigma(x) = x - \pi(x) = x - x' = x''$.

(3.2) $P + Q = E_n$.

(3.3) $PQ = QP = 0$.

(4) *P ist positiv mit $\sigma(P) \subseteq \{0, 1\}$ und* $\mathrm{tr}(P) = \mathrm{rg}(P)$.

(5) *Sei X nicht der Nullraum, $\mathcal{V} = \{v_1, \ldots, v_r\}$ eine Basis von X und*

$$V = (v_1, \ldots, v_r) \in \mathbb{R}^{n \times r}$$

die Matrix mit den Spalten $v_1, \ldots, v_r$. Dann ist die quadratische Matrix

$$V^t V \in \mathbb{R}^{r \times r}$$

invertierbar. Es gilt die Projektionsformel

$$P = V(V^t V)^{-1} V^t\,.$$

(6) *Wenn in* (5) *die Vektoren $v_1, \ldots, v_r$ paarweise orthogonal sind, vereinfacht sich die Projektionsformel zu*

$$P = \sum_{\rho=1}^{r} \frac{v_\rho \otimes v_\rho}{\|v_\rho\|^2} = \sum_{\rho=1}^{r} \frac{v_\rho v_\rho^t}{\|v_\rho\|^2}\,.$$

Siehe Satz und Definition 5.5.

Beweis. Nachweis von (1.3) *und* (1.4). Als Durchschnitt der Kerne der linearen Abbildungen $x \mapsto \langle x, v \rangle$ mit $v \in X$ ist $X^\perp$ ein Teilvektorraum des $\mathbb{R}^n$. Analog folgt, dass $X^{\perp\perp}$ ein Teilvektorraum des $\mathbb{R}^n$ ist. Die Aussage (1.4) folgt aus (1.3).

Die Aussage (1.3) gilt in den Fällen $X = \{0\}$ und $X = \mathbb{R}^n$ trivialerweise. Sei also $1 \le r = \dim X \le n - 1$ mit $n \ge 2$.

Wir wählen eine Basis $\{a_1, \ldots, a_r\}$ von X. Sei $A = (a_1, \ldots, a_r) \in \mathbb{R}^{n \times r}$. Dann gelten

$$X = \mathrm{im}(A)\,, \quad X^\perp = \ker(A^t)\,, \quad 1 \le \dim(X^\perp) = n - r \le n - 1\,.$$

Dabei verwenden wir die Beziehung $\mathrm{rg}(A^t) = \mathrm{rg}(A) = r$ und die Dimensionsformel für die transponierte Matrix $A^t \in \mathbb{R}^{r\times n}$. Siehe Satz 6.10. Wir wählen eine Basis $\{a_{r+1}, \ldots a_n\}$ des Orthogonalraumes $X^\perp$. Dann ist die Menge $\{a_1, \ldots, a_n\}$ eine Basis des $\mathbb{R}^n$.

Sei $x \in \mathbb{R}^n$. Dann gibt es eindeutig bestimmte $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit

$$x = \alpha_1 a_1 + \ldots + \alpha_n a_n \,.$$

Die Vektoren

$$x' = \alpha_1 a_1 + \ldots + \alpha_r a_r \in X \,, \quad x'' = \alpha_{r+1} a_{r+1} + \ldots + \alpha_n a_n \in X^\perp$$

bilden eine orthogonale Zerlegung $x = x' + x''$. Seien $x_1 \in X$ und $x_2 \in X^\perp$ irgendwelche Vektoren mit $x = x_1 + x_2$. Dann gibt es eindeutig bestimmte Koeffizienten $\beta_1, \ldots, \beta_n \in \mathbb{R}$ mit

$$x_1 = \beta_1 a_1 + \ldots + \beta_r a_r \,, \quad x_2 = \beta_{r+1} a_{r+1} + \ldots + \beta_n a_n \,.$$

Weil $\{a_1, \ldots, a_n\}$ eine Basis des $\mathbb{R}^n$ ist, folgt $a_\nu = \beta_\nu$ für alle $\nu = 1, \ldots, n$. Daher ist eine orthogonale Zerlegung eines beliebigen Vektors des $\mathbb{R}^n$ in einen Vektor aus X und einen Vektor aus $X^\perp$ eindeutig bestimmt. Diese Zerlegung hängt insbesondere nicht von der Wahl der Basen der Teilräume X und $X^\perp$ ab. Damit ist (1.3) bewiesen.

Die Gültigkeit der Aussagen in (2) und (3) ist klar.

Nachweis von (4). Seien $\lambda \in \mathbb{R}$ und $v \in \mathbb{R}^n \setminus \{0\}$ mit $Pv = \lambda v$ gegeben. Aus

$$\lambda v = Pv = P^2 v = P(\lambda v) = \lambda^2 v$$

folgt $\lambda^2 = \lambda$. Also gilt $\sigma(P) \subseteq \{0, 1\}$.

Offenbar gilt $\mathrm{tr}(P) = \mathrm{rg}(P)$, wenn $P = 0$ oder $P = E_n$ gilt. Wir betrachten den Fall $1 \le r = \dim X = \mathrm{im}(P) \le n - 1$. Sei eine Basis $\{a_1, \ldots, a_r\}$ von X und $\{a_{r+1}, \ldots a_n\}$ eine Basis von $X^\perp$. Dann ist $\{a_1, \ldots, a_n\}$ eine Basis des $\mathbb{R}^n$ aus Eigenvektoren der Projektionsmatrix P. Sei $W = (a_1, \ldots, a_n) \in \mathrm{GL}(n, \mathbb{R})$. Dann gilt

$$W^{-1} P W = \begin{pmatrix} E_r & 0_{r\times(n-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(n-r)} \end{pmatrix} .$$

Mit Satz erhalten wir

$$\mathrm{tr}(P) = \mathrm{tr}(W^{-1} P W) = \mathrm{tr}(E_r) = r = \dim(X) = \mathrm{rg}(P) \,.$$

Damit ist (4) bewiesen.

Nachweis von (5). Zuerst zeigen wir, dass die $(r \times r)$-Matrix $V^t V$ invertierbar ist. Weil die Vektoren $v_1, \ldots, v_r$ linear unabhängig sind, gilt $\mathrm{rg}(V) = r$. Sei $y \in \mathbb{R}^r$ mit $V^t V y = 0$ gegeben. Dann gilt

$$\langle Vy, Vy \rangle = \langle y, V^t V y \rangle = 0 \,.$$

Also gilt $Vy = 0$. Weil die Vektoren $v_1, \dots, v_r$ linear unabhängig sind, folgt $y = 0$. Also ist die quadratische Matrix V^tV invertierbar.

Nun sei $x \in \mathbb{R}^n$ beliebig gewählt. Dann gilt $Px \in X = \mathrm{im}(P)$. Weil $\mathcal{V}$ eine Basis von X ist, gibt es $\xi_1, \dots, \xi_r \in \mathbb{R}$ mit

$$Px = \xi_1 v_1 + \ldots + \xi_r v_r \, .$$

Wir betrachten den Vektor

$$\xi = (\xi_\rho)_{\rho=1,\dots,r} = \xi_1 e_1^{(r)} + \ldots + \xi_r e_r^{(r)} \in \mathbb{R}^r \, .$$

Dabei ist $\mathcal{E}_r = \{e_1^{(r)}, \dots, e_n^{(r)}\}$ die kanonische Basis des $\mathbb{R}^r$. Nach Konstruktion gilt

$$Px = V\xi \in \mathbb{R}^n \, .$$

Wegen $x - Px \in X^\perp$ gilt

$$v_\rho^t(x - V\xi) = \langle v_\rho, x - V\xi \rangle = \langle v_\rho, x - Px \rangle = 0$$

für alle $\rho = 1, \dots, r$. Also ist der Vektor $\xi \in \mathbb{R}^r$ eine Lösung der *Normalgleichung*

$$V^tV\xi = V^tx \in \mathbb{R}^r \, .$$

Weil $V^tV \in \mathbb{R}^{r\times r}$ invertierbar ist, folgt

$$\xi = (V^tV)^{-1}V^tx \, .$$

Multiplikation mit V ergibt

$$Px = V\xi = V(V^tV)^{-1}V^tx \, .$$

Damit ist die Projektionsformel

$$P = V(V^tV)^{-1}V^t$$

bewiesen.

Nachweis von (6). Wenn $v_1, \dots, v_r$ paarweise orthogonal sind, gilt

$$V^tV = \mathrm{diag}(\|v_1\|^2, \dots, \|v_r\|^2) \, .$$

Die Projektionsformel liefert in diesem Fall

$$Px = V\left(\frac{v_1^tx}{\|v_1\|^2}, \dots, \frac{v_r^tx}{\|v_r\|^2}\right)^t = \frac{v_1^tx}{\|v_1\|^2}\, v_1 + \ldots + \frac{v_r^tx}{\|v_r\|^2}\, v_r$$

für alle $x \in \mathbb{R}^n$. Damit ist der Beweis beendet. □

Beispiel 16.3. Sei $X \subseteq \mathbb{R}^3$ der Teilvektorraum, der von den beiden Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

erzeugt wird. Sei $P \in \mathbb{R}^{3\times 3}$ die Matrixdarstellung der orthonalen Projektion des $\mathbb{R}^3$ auf X bezüglich der kanonischen Basis. Dann gilt

$$\begin{aligned} P = V(V^tV)^{-1}V^t &= \begin{pmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \circ \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}^{-1} \circ \begin{pmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \circ \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{pmatrix}. \end{aligned}$$

Siehe Beispiel 16.5. □

Satz und Definition 16.4. *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\, .\,,.\, \rangle$ und der euklidischen Norm $\|\,.\,\|$ versehen. Sei $X \subseteq \mathbb{R}^n$ ein Teilvektorraum mit $\dim(X) = r \in \mathbb{N}$.*

(1) *Eine Basis $\{v_1, \ldots, v_r\}$ von X heißt eine* Orthogonalbasis *von X, wenn $\langle v_\rho, v_\sigma \rangle = 0$ für alle $\rho, \sigma = 1, \ldots, r$ mit $\rho \neq \sigma$ gilt.*

(2) *Eine Basis $\{v_1, \ldots, v_r\}$ von X heißt eine* Orthonormalbasis *von X, wenn $\langle v_\rho, v_\sigma \rangle = \delta_{\rho\sigma}$ für alle $\rho, \sigma = 1, \ldots, r$ gilt.*

(3) *Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt eine* Orthogonalmatrix, *wenn $A^tA = E_n$ gilt.*

(4) *Die Menge der Orthogonalmatrizen $A \in \mathbb{R}^{n\times n}$ bezeichnen wir mit $\mathrm{O}(n, \mathbb{R})$.*

Es gelten die folgenden Aussagen (5), (6), (7).

(5) *Die kanonische Basis $\mathcal{E}_n$ ist eine Orthonormalbasis des $\mathbb{R}^n$.*

(6) *Sei $\{v_1, \ldots, v_r\}$ eine Basis von X. Dann liefert das folgende induktive Verfahren von* Gram-Schmidt *eine Orthogonalbasis $\{w_1, \ldots, w_r\}$ von X.*

(6.1) *Wir setzen $w_1 = v_1$.*

(6.2) *Wir berechnen*

$$w_k = v_k - \sum_{\rho=1}^{k-1} \frac{\langle v_k, w_\rho \rangle}{\langle w_\rho, w_\rho \rangle}\, w_\rho\,,$$

sukzessive für $k = 2, \ldots, n$.

(7) *Normieren der Vektoren einer Orthogonalbasis von X ergibt eine Orthonormalbasis von X.*

(8) *Eine Matrix $A = (a_1, \dots, a_n) \in \mathbb{R}^{n\times n}$ ist genau dann eine Orthogonalmatrix, wenn ihre Spalten $a_1, \dots, a_n$ eine Orthonormalbasis des $\mathbb{R}^n$ bilden.*

Beispiel 16.5 (Fortsetzung von 16.3)**.** Wir wenden das Verfahren von Gram-Schmidt auf die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} , \quad v_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

an. Anschließend berechnen wir bezüglich der kanonischen Basis $\mathcal{E}_3$ mit Hilfe der orthogonalen Vektoren w_1, w_2 die Matrixdarstellung P der orthogonalen Projektion auf $X = [v_1, v_2] = [w_1, w_2]$.

$$w_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} , \quad w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle}\, w_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} .$$

Mit Formel (8) aus dem Projektionssatz 16.2 erhalten wir

$$P = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{pmatrix}$$

Siehe Beispiel 16.3. □

Nach diesen Vorbereitungen können wir uns dem Spektralsatz zuwenden.

Satz 16.6 (Spektralsatz. Reelle Version)**.** *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\, .\, ,\, .\, \rangle$ und der euklidischen Norm $\|\, .\, \|$ versehen. Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann gelten* (1) *bis* (9)*.*

(1) *$\chi_A(t)$ zerfällt vollständig über $\mathbb{R}$, das heißt, es gibt $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit*

$$\chi_A(t) = (-1)^n \prod_{k=1}^{n} (t - \lambda_k) .$$

(2) *Eigenvektoren von A, die zu verschiedenen Eigenwerten von A gehören, sind zueinander orthogonal.*

(3) *Sei $v \in \mathbb{R}^n \setminus \{0\}$ ein Eigenvektor von A. Wenn $x \in \mathbb{R}^n$ orthogonal zu v ist, dann ist auch Ax orthogonal zu v.*

(4) *Der* $\mathbb{R}^n$ *besitzt eine Orthonormalbasis* $\mathcal{V} = \{v_1, \ldots, v_n\}$ *aus Eigenvektoren der symmetrischen Matrix* A*. Sei* $V = (v_1, \ldots, v_n)$ *die Matrix mit den Spalten* $v_1, \ldots, v_n$*. Dann gilt*

$$\operatorname{diag}(\lambda_1, \ldots, \lambda_n) = V^{-1} \circ A \circ V\,.$$

Die Matrix V *ist eine Orthogonalmatrix.*

(5) A *ist diagonalisierbar über* $\mathbb{R}$.

(6) *Bezüglich der kanonischen Basis des* $\mathbb{R}^n$ *sei die Projektionsmatrix* P_λ *die Matrixdarstellung der orthogonalen Projektion auf den Eigenraum*

$$\operatorname{im}(P_\lambda) = \ker(A - \lambda E_n)$$

zum Eigenwert $\lambda \in \sigma(A)$*. Dann besitzt die Einheitsmatrix* $E_n \in \mathbb{R}^{n \times n}$ *die Darstellung*

$$E_n = \sum_{\lambda \in \sigma(A)} P_\lambda\,,$$

wobei $P_\lambda P_{\lambda'} = 0$ *für alle* $\lambda,\ \lambda' \in \sigma(A)$ *mit* $\lambda \neq \lambda'$ *gilt.*

(7) *Die symmetrische Matrix* A *besitzt die* Spektralzerlegung

$$A = \sum_{\lambda \in \sigma(A)} \lambda\, P_\lambda\,.$$

Die Projektionsmatrizen $P_\lambda \in \mathbb{R}^{n \times n}$ *mit* $P_\lambda P_{\lambda'} = 0$ *für alle* $\lambda,\ \lambda' \in \sigma(A)$ *mit* $\lambda \neq \lambda'$ *sind eindeutig durch* A *bestimmt.*

(8) *Für alle Polynome* $f(t) \in \mathbb{R}[t]$ *gilt*

$$f(A) = \sum_{\lambda \in \sigma(A)} f(\lambda)\, P_\lambda\,.$$

(9) A *ist genau dann invertierbar, wenn* $0 \notin \sigma(A)$ *gilt. In diesem Fall gilt*

$$A^{-1} = \sum_{\lambda \in \sigma(A)} \lambda^{-1} P_\lambda\,.$$

(10) *Die symmetrische Matrix* A *ist genau dann positiv, wenn* $\sigma(A) \subseteq [0, \infty)$ *gilt. In diesem Fall gibt es eine eindeutig bestimmte positive Matrix* $B \in \mathbb{R}^{n \times n}$ *mit* $A = B^2$*. Diese Matrix* B *heißt die* positive Quadratwurzel $A^{\frac{1}{2}}$ *von* A*. Es gilt*

$$A^{\frac{1}{2}} = \sum_{\lambda \in \sigma(A)} \lambda^{\frac{1}{2}}\, P_\lambda\,.$$

Beweis. Zum Nachweis von Aussage (1) betrachten wir die symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ als ein Element von $\mathbb{C}^{n \times n}$. Das charakteristische Polynom

$$\chi_A(\tau) = \det(A - \tau E_n)$$

betrachten wir als ein Element von $\mathbb{C}[\tau]$. Um Verwechslungen mit der Transposition auszuschließen, bezeichnen wir die Unbekannte mit τ.

Nach dem Fundamentalssatz 15.12 der Algebra zerfällt $\chi_A(\tau)$ über $\mathbb{C}$ vollständig in Linearfaktoren, das heißt, es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit

$$\chi_A(\tau) = (-1)^n \prod_{k=1}^{n} (\tau - \lambda_k) .$$

Dabei ist jede Nullstelle entsprechend ihrer algebraischen Vielfachheit mehrfach aufgeführt. Wir zeigen, dass $\lambda_1, \ldots, \lambda_n$ reell sind.

Folgende Schreibweise ist üblich: Der Vektor $\overline{z} \in \mathbb{C}^n$ und die Matrix $\overline{C} \in \mathbb{C}^{n \times n}$ entstehen aus $z = (z_\nu) \in \mathbb{C}^n$ und $C = (c_{\nu\mu}) \in \mathbb{C}^{n \times n}$, indem die Komponenten z_ν und die Einträge $c_{\nu\mu}$ durch die konjugiert-komplexen Zahlen $\overline{z}_\nu$ respektive $\overline{c}_{\nu\mu}$ ersetzt werden. Weil A eine reelle Matrix ist, gilt $\overline{A} = A$.

Sei $\lambda \in \mathbb{C}$ eine beliebige Nullstelle von $\chi_A(\tau)$. Dann gibt es zu λ einen Eigenvektor $v = (v_\nu) \in \mathbb{C}^n \setminus \{0\}$ von A. Es gilt

$$\overline{v}^t v = \sum_{\nu=1}^{n} |v_\nu|^2 \neq 0 , \quad \overline{v} \in \mathbb{C}^n \setminus \{0\} .$$

Über dem Körper $\mathbb{C}$ gelten die Gleichungen

$$Av = \lambda v , \quad A\overline{v} = \overline{Av} = \overline{\lambda}\,\overline{v} .$$

Die zweite Gleichung folgt aus der ersten, weil A reell ist. Nun folgen die beiden Matrixgleichungen

$$\overline{v}^t A v = \lambda\, \overline{v}^t v , \quad v^t A \overline{v} = \overline{\lambda}\, v^t \overline{v} .$$

Wir transponieren die zweite Matrixgleichung und erhalten

$$(\overline{v}^t A) v = \overline{\lambda}\, \overline{v}^t v ,$$

weil A symmetrisch ist. Es folgt

$$0 = \overline{v}^t A v - \overline{v}^t A v = (\lambda - \overline{\lambda})\, \overline{v}^t v .$$

Wegen $\overline{v}^t v \neq 0$ folgt $\lambda \in \mathbb{R}$. Damit ist Aussage (1) bewiesen.

Wir beweisen (2). Seien $v, w \in \mathbb{R}^n \setminus \{0\}$ Eigenvektoren zu A zu verschiedenen Eigenwerten $\lambda, \mu \in \mathbb{R}$. Weil $A \in \mathbb{R}^{n \times n}$ symmetrisch ist, gilt

$$\lambda \langle v, w \rangle = \langle Av, w \rangle = \langle v, Aw \rangle = \mu \langle v, w \rangle .$$

Aus $\lambda \neq \mu$ folgt $\langle v, w \rangle = 0$. Damit ist (2) bewiesen.

Nachweis von (3). Sei $v \in \mathbb{R}^n \setminus \{0\}$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{R}$. Wenn $x \in \mathbb{R}^n$ orthogonal zu v ist, dann ist auch Ax orthogonal zu v. Es gilt nämlich $\langle Ax, v \rangle = \langle x, Av \rangle = \lambda \langle x, v \rangle = 0$.

Nachweis von (4). Sei $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ die Abbildung mit

$$\varphi(x) = Ax\,, \quad M_{\mathcal{E}\mathcal{E}}(\varphi) = A\,.$$

Dabei ist $\mathcal{E} = \{e_1, \ldots, e_n\}$ die kanonische Basis von $\mathbb{R}^n$. Wir zeigen durch vollständige Induktion nach n, dass eine orthogonale Matrix $V \in \mathbb{R}^{n\times n}$ mit

$$V^t \circ A \circ V = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

existiert. Im Fall $n = 1$ ist die Aussage mit $V = E_1$ trivialerweise erfüllt.

Sei $k \in \mathbb{N}$ gegeben. Wir nehmen an, dass die Aussage für alle symmetrischen Matrizen aus $\mathbb{R}^{k\times k}$ gilt. Sei eine beliebige $A \in \mathbb{R}^{n\times n}$ symmetrische Matrix, wobei $n = k + 1$ gilt.

Zum Eigenwert $\lambda_1 \in \mathbb{R}$ gibt es einen normierten Eigenvektor $v_1 \in \mathbb{R}^n$. Der Eigenvektor v_1 kann zu einer Basis des $\mathbb{R}^n$ ergänzt werden. Aus dieser Basis konstruieren wir mit dem Verfahren von Gram-Schmidt eine Orthogonalbasis. Der erste Vektor dieser Basis ist der Vektor v_1. Normieren liefert eine Orthonormalbasis $\mathcal{C} = \{v_1, c_2, \ldots, c_n\}$. Sei

$$C = (v_1, c_2, \ldots, c_n) \in \mathrm{O}(n, \mathbb{R})\,, \quad C^{-1} = C^t$$

die entsprechende Orthogonalmatrix mit den Spalten $v_1, c_2, \ldots, c_n$.

Es gilt $\varphi(v_1) = \lambda_1 v_1$. Nach Aussage (3) bildet die lineare Abbildung φ den Teilvektorraum $[c_2, \ldots, c_n] \subseteq \mathbb{R}^n$ in sich ab. Sei

$$A_1 = (\alpha_{\nu\mu})_{\nu,\mu=2,\ldots,n} \in \mathbb{R}^{k\times k}$$

die Matrix, deren Einträge $\alpha_{\nu\mu}$ durch die Gleichungen

$$\varphi(c_\nu) = \sum_{\mu=2}^{n} \alpha_{\mu\nu} c_\mu\,, \quad \nu = 2, \ldots n$$

eindeutig bestimmt sind. Siehe Satz und Definition 7.4. Dann gilt

$$M_{\mathcal{C}\mathcal{C}}(\varphi) = C^t A C = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix}.$$

Weil A symmetrisch ist, sind $C^t A C$ und A_1 symmetrisch. Wegen

$$\chi_A(\tau) = (\lambda_1 - \tau) \cdot \chi_{A_1}(\tau)$$

sind $\lambda_2, \ldots, \lambda_n$ die Eigenwerte der symmetrischen Matrix $A_1 \in \mathbb{R}^{k\times k}$. Nach Induktionsvoraussetzung gibt es eine Matrix $C_1 \in \mathrm{O}(k, \mathbb{R})$ mit

$$C_1^t A_1 C_1 = \operatorname{diag}(\lambda_2, \ldots, \lambda_n)\,.$$

Folglich gilt

$$\begin{pmatrix} 1 & 0 \\ 0 & C_1 \end{pmatrix}^t \circ \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 0 & C_1 \end{pmatrix} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)\,.$$

Nach Konstruktion ist

$$V = C \circ \begin{pmatrix} 1 & 0 \\ 0 & C_1 \end{pmatrix} \in \mathrm{O}(n, \mathbb{R})$$

eine Orthogonalmatrix mit

$$V^t \circ A \circ V = \operatorname{diag}(\lambda_1, \dots, \lambda_n)\,.$$

Die erste Spalte von V ist der eingangs betrachtete Eigenvektor v_1. Die anderen Spalten bezeichnen wir der Reihe nach mit $v_2, \dots, v_n$. Die Vektoren $v_1, \dots, v_n$ sind normierte Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_n$. Außerdem ist $\mathcal{V} = \{v_1, \dots, v_n\}$ eine Orthonormalbasis des $\mathbb{R}^n$. Damit sind (4) und (5) bewiesen.

Die Aussagen (6) bis (10) sind Konsequenzen der Aussagen (1) bis (5). Wir beweisen exemplarisch die Existenz und die Einzigkeit der positiven Quadratwurzel einer positiven Matrix A. Es gilt $\sigma(A) \subseteq [0, \infty)$. Also ist

$$B = \sum_{\lambda \in \sigma(A)} \lambda^{\frac{1}{2}} P_\lambda \in \mathbb{R}^{n \times n}$$

eine positive Matrix mit $A = B^2$. Sei $C \in \mathbb{R}^{n \times n}$ eine positive Matrix mit $A = C^2$. Dann gibt es eine Orthonormalbasis $\mathcal{W} = \{w_1, \dots, w_n\}$ des $\mathbb{R}^n$ aus Eigenvektoren von C. Sei $w \in \mathcal{W}$. Dann gibt es $\mu \in \mathbb{R}$ mit $Cw = \mu w$. Es folgt

$$Aw = C(Cw) = \mu^2 w\,, \quad \mu^2 \in \sigma(A)\,, \quad w \in \operatorname{im}(P_{\mu^2})\,.$$

Demnach ist der definierende Ausdruck der Matrix B die Spektralzerlegung der Matrix C. Also gilt $B = C$. Damit ist der Beweis beendet. □

Beispiel 16.7. Der Vektorraum $\mathbb{R}^3$ sei mit dem euklidischen inneren Produkt $\langle\, .\, ,\, .\, \rangle$ und der euklidischen Norm $\|\, .\, \|$ versehen. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $v_1,\, v_2,\, v_3 \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} -\frac{3}{2} & -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{3}{2} & -1 \\ 1 & -1 & 0 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}.$$

Die Vektoren v_1, v_2, v_3 bilden eine Basis $\mathcal{V}$ des $\mathbb{R}^3$ aus Eigenvektoren von A. Sei λ_k der Eigenwert zu v_k. Es gilt

$$\lambda_1 = \lambda_2 = -2\,, \quad \lambda_3 = 1\,.$$

Mit 16.3 und 16.5 erhalten wir

$$P_{-2} = \frac{1}{6}\begin{pmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{pmatrix}, \quad P_1 = E_3 - P_{-2} = \frac{1}{6}\begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{pmatrix}.$$

Die Matrix A besitzt die Spektralzerlegung

$$A = (-2) \cdot P_{-2} + P_1 = -\frac{2}{6} \begin{pmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{pmatrix} .$$

Wir berechnen A^{10} und A^{-1} mit dem Spektralkalkül.

$$A^{10} = (-2)^{10} P_{-2} + P_1 = \frac{1}{2} \begin{pmatrix} 1707 & 341 & -682 \\ 341 & 1707 & 682 \\ -682 & 682 & 684 \end{pmatrix} ,$$

$$A^{-1} = (-2)^{-1} P_{-2} + P_1 = \frac{1}{4} \begin{pmatrix} -1 & -1 & 2 \\ -1 & -1 & -2 \\ 2 & -2 & 2 \end{pmatrix} .$$

□

17 Ergänzung: Kleinste Fehlerquadrate. Gauß-Legendre

Die Methode der kleinsten Fehlerquadrate geht auf Gauß und Legendre zurück. Wir betrachten in dieser Ergänzung nur die einfachsten Situationen.

In der (X, Y)-Ebene $\mathbb{R}^2$ seien $n \geq 2$ Punkte

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \dots, \begin{pmatrix} a_n \\ b_n \end{pmatrix} \tag{17.1}$$

gegeben. Dabei setzen wir voraus, dass die Komponenten $a_1, \dots, a_n$ paarweise verschieden sind. Wir suchen eine Gerade

$$Y = \xi_1 + \xi_2 X \tag{17.2}$$

in der (X, Y)-Ebene derart, dass die Summe

$$\sum_{\nu=1}^{n} \big((\xi_1 + \xi_2 a_\nu) - b_\nu\big)^2 \tag{17.3}$$

der Quadrate der vertikalen Fehler minimal wird. Die Analysis liefert für die unbekannten Koeffizienten ξ_1, $\xi_2 \in \mathbb{R}$ das lineare Gleichungsystem

$$\begin{pmatrix} n & \sum\limits_{\nu=1}^{n} a_\nu \\ \sum\limits_{\nu=1}^{n} a_\nu & \sum\limits_{\nu=1}^{n} a_\nu^2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \sum\limits_{\nu=1}^{n} b_\nu \\ \sum\limits_{\nu=1}^{n} a_\nu b_\nu \end{pmatrix}. \tag{17.4}$$

Die (2×2)-Matrix auf der linken Seite von (17.4) ist nach Voraussetzung invertierbar. Also besitzt (17.4) genau eine Lösung $(\xi_1, \xi_2)^t \in \mathbb{R}^2$.

Wir untersuchen, durch welche algebraische Umformung das Gleichungssystem (17.4) aus (17.1) und (17.2) hervorgeht. Zunächst ist klar, dass die n Punkte die Gleichung (17.2) erfüllen *müssten*. Wir fassen diese n Gleichungen in der Matrixgleichung

$$\begin{pmatrix} 1 & a_1 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \tag{17.5}$$

zusammen. Aber das Gleichungssystem (17.5) enthält im Fall $n > 2$ mehr Gleichungen als Unbekannte und besitzt daher im Allgemeinen keine Lösung. Wir transponieren die $(n \times 2)$-Matrix auf der linken Seite. Mit der transponierten $(2 \times n)$-Matrix multiplizieren wir das Gleichungssystem (17.5) und erhalten

$$\begin{pmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_n \end{pmatrix} \begin{pmatrix} 1 & a_1 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \tag{17.6}$$

Dies ist etwas anders aufgeschrieben das eindeutig lösbare lineare Gleichungssystem (17.4). Die Matrixgleichung (17.6) ist die sogenannte *Normalgleichung* der Matrixgleichung (17.5).

Beispiel 17.1. In der (X, Y)-Ebene $\mathbb{R}^2$ seien die Punkte

$$\begin{pmatrix} -2 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

gegeben. Bestimmen Sie nach Methode der kleinsten Fehlerquadrate die zugehörige Ausgleichsgerade $Y = \xi_1 + \xi_2 X$.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ -1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \end{pmatrix}.$$

$$\begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \end{pmatrix}.$$

$$\begin{aligned} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} &= \begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 8 \end{pmatrix} \\ &= \frac{1}{20} \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} \\ &= \frac{1}{20} \begin{pmatrix} 28 \\ 36 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 7 \\ 9 \end{pmatrix}. \end{aligned}$$

$$Y = \tfrac{7}{5} + \tfrac{9}{5} X.$$

Wir beschreiben die geometrische Bedeutung der Normalgleichung. Dafür ändern wir (17.5) leicht ab. Statt der $(n \times 2)$-Matrix

$$\begin{pmatrix} 1 & a_1 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix}$$

betrachten wir eine beliebige Matrix $A \in \mathbb{R}^{n \times m}$ mit $n \geq m \geq 2$ und $\mathrm{rg}(A) = m$. Weil A maximalen Rang besitzt, enthält der Kern von A nur den Nullvektor. Die quadratische Matrix $A^t A$ ist daher invertierbar. Der eventuell unlösbaren Gleichung

$$Ax = b \tag{17.7}$$

mit gegebenem Vektor $b \in \mathbb{R}^n$ ordnen wir die eindeutig lösbare *Normalgleichung*

$$A^t Ax = A^t b \tag{17.8}$$

zu. In den Anwendungen liefern die erhobenen Daten die Einträge der Matrix A und des Vektors b. Die eindeutig bestimmte Lösung ξ der Normalgleichung liefert die Koeffizienten des gesuchten Ausgleichsobjektes. Beispielsweise kann ein Polynom vom Grad $m-1$ das Ausgleichsobjekt sein. Siehe Beispiel 17.3.

Die geometrische Interpretation der Normalgleichung beruht auf dem Umstand, dass die Matrix

$$P = A \circ (A^t A)^{-1} \circ A^t$$

eine Projektionsmatrix mit demselben Bild wie die Matrix A ist. Das Lösen der Normalgleichung kann unter den obigen Voraussetzungen durch das Lösen der Gleichung

$$Ax = Pb \tag{17.9}$$

ersetzt werden. Durch Abänderung der Inhomogenität b in den Fußpunkt Pb wird die Gleichung (17.7) lösbar gemacht. Wenn (17.7) lösbar ist, dann gilt $b \in \mathrm{im}(A)$. In diesem Fall stimmen die beiden Gleichungen (17.9) und (17.7) überein.

Satz 17.2. *Seien $m, n \in \mathbb{N}$ mit $2 \leq m \leq n$. Es seien $A \in \mathbb{R}^{n \times m}$ mit $\mathrm{rg}(A) = m$ und $b \in \mathbb{R}^m$ beliebig gegeben. Dann gelten die folgenden Aussagen.*

(1) $A^t A \in \mathrm{GL}(n, \mathbb{R}^n)$.

(2) $P = A(A^t A)^{-1} A^t \in \mathrm{Proj}(\mathbb{R}^{n \times n})$.

(3) $\mathrm{im}(P) = \mathrm{im}(A)$.

(4) *Die Normalgleichung*

$$A^t Ax = A^t b$$

besitzt genau eine Lösung $\xi \in \mathbb{R}^m$.

(5) *Die Lösung* ξ *der Normalgleichung ist die einzige Lösung von*

$$Ax = Pb\,.$$

Dies folgt aus (3) *und* $\ker(A) = \{0\}$.

(6) *Extremaleigenschaft. Es gilt*

$$\|A\xi - b\| = \min_{x\in\mathbb{R}^m} \|Ax - b\|\,.$$

Dies folgt aus (2), (3), (5) *und dem Satz des Pythagoras. Dabei ist* $\|\,.\,\|$ *das euklidische innere Produkt des* $\mathbb{R}^n$.

(7) *Die Lösung* ξ *der Normalgleichung ist eine Lösung des Extremalproblems, die euklidische Norm des Vektors* $Ax - b$ *möglichst klein zu machen.*

Beispiel 17.3. In der Ebene $\mathbb{R}^2$ seien die Punkte

$$\begin{pmatrix} -2 \\ 8 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} 2 \\ \frac{5}{2} \end{pmatrix}$$

gegeben. Bestimmen Sie nach Methode der kleinsten Fehlerquadrate die zugehörige Ausgleichsparabel $Y = \alpha + \beta X + \gamma X^2$.

$$\begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \\ x_1^2 & \dots & x_n^2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \\ x_1^2 & \dots & x_n^2 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}.$$

$$\begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}^{-1} = \frac{1}{70} \begin{pmatrix} 34 & 0 & -10 \\ 0 & 7 & 0 \\ -10 & 0 & 5 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 8 \\ 2 \\ 1 \\ \frac{1}{2} \\ \frac{5}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 16 \\ 4 \\ 2 \\ 1 \\ 5 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 28 \\ -25 \\ 89 \end{pmatrix}.$$

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \frac{1}{2 \cdot 70} \begin{pmatrix} 34 & 0 & -10 \\ 0 & 7 & 0 \\ -10 & 0 & 5 \end{pmatrix} \begin{pmatrix} 28 \\ -25 \\ 89 \end{pmatrix} = \frac{1}{140} \begin{pmatrix} 62 \\ -175 \\ 165 \end{pmatrix}.$$

Ausgleichsparabel.

$$Y = \tfrac{31}{70} - \tfrac{5}{4}\, X + \tfrac{33}{28}\, X^2 .$$

Dezimalentwicklung der Koeffizienten des Polynoms auf zehn Stellen.

$$Y = 0.4428571429 - 1.250000000\, X + 1.178571429\, X^2 .$$

18 Ergänzung: Cholesky-Zerlegung

In diesem Abschnitt betrachten wir die strikte Bruhat-Zerlegung

$$A = LDPU$$

einer beliebigen positiv definiten reellen Matrix A. Weil die Hauptminoren von A positiv sind, gilt $P = E$. Wegen $A = A^t$ gilt $U = L^t$. Daher besitzt A die strikte Bruhat-Zerlegung

$$A = LDL^t \,.$$

Eine solche Zerlegung heißt eine *Cholesky-Zerlegung* von A. Aus der Einzigkeit der strikten Bruhat-Zerlegung folgt die Einzigkeit der Cholesky-Zerlegung. Das Verfahren zur Berechnung der strikten Bruhat-Zerlegung liefert im Fall einer positiv definiten Matrix A die Cholesky-Zerlegung. Die Diagonalelemente von D sind positiv.

Definition 18.1. *Seien $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n\times n}$ gegeben. Für $k = 1, \ldots, n$ sei*

$$A_{1:k,1:k} = (a_{\nu\mu})_{\nu,\mu=1,\ldots,k}$$

die linke obere $(k \times k)$-Teilmatrix, die aus A durch Streichen der letzten $(n-k)$ Zeilen und Spalten entsteht. Die Determinante

$$\delta_k(A) = \det(A_{1:k,1:k})$$

heißt der k-te Hauptminor *von A. Offenbar gilt $\delta_n(A) = \det(A)$.*

Eine symmetrische reelle Matrix ist genau dann positiv definit, wenn alle ihre Hauptminoren positiv sind.

Satz und Definition 18.2. *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\, .\,,.\, \rangle$ und der euklidischen Norm $\|\,.\,\|$ versehen. Sei eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$. Dann sind die Aussagen (1) bis (6) äquivalent.*

(1) *Es gilt $\langle Ax, x\rangle > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.*

(2) *Es gibt $B \in \mathrm{GL}(n, \mathbb{R})$ mit $A = B^t B$.*

(3) *A ist invertierbar und positiv.*

(4) *Alle Eigenwerte $\lambda_1, \ldots, \lambda_n$ von A sind positiv.*

(5) *Alle Hauptminoren $\delta_1(A), \ldots, \delta_n(A)$ von A sind positiv.*

Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$, die eine der Bedingungen (1) bis (6) erfüllt, heißt positiv definit.

Nach 18.2 ist eine Diagonalmatrix genau dann positiv, wenn alle Diagonalelemente positiv sind. Die Aussage (2) in 18.2 lässt sich verschärfen. Die Matrix B kann als positiv definite obere Dreiecksmatrix gewählt werden.

Satz und Definition 18.3 (Cholesky-Zerlegung)**.** *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\, .\, ,\, .\, \rangle$ und der euklidischen Norm $\|\, .\, \|$ versehen. Sei $A = (a_{\nu\mu}) \in \mathbb{R}^{n\times n}$ eine positiv definite Matrix. Dann gelten folgende Aussagen und Definitionen.*

(1) *Es gibt eine normierte untere Dreiecksmatrix $L \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$ und eine positiv definite Diagonalmatrix $D \in \mathrm{D}^{\times}(n, \mathbb{R})$ mit*

$$A = LDL^t\,.$$

Die Matrizen $L \in \mathrm{L}_1^{\llcorner}(n, \mathbb{R})$ und $D \in \mathrm{D}^{\times}(n, \mathbb{R})$ sind eindeutig bestimmt.

(2) *Die obige Zerlegung $A = LDL^t$ heißt die* Sandwichform der Cholesky-Zerlegung *der positiv definiten Matrix A.*

(3) *Die Cholesky-Zerlegung von A ist die strikte Bruhat-Zerlegung $A = LDPU$ mit $P = E_n$ und $U = L^t$.*

(4) *Seien $\delta_1(A), \dots, \delta_n(A)$ die Hauptminoren von A und $d_1(A), \dots, d_n(A)$ die Diagonalelemente von D. Dann gilt*

$$\delta_k(A) = \prod_{\nu=1}^{n} d_\nu(A)$$

für $k = 1, \dots, n$. Insbesondere gelten die Beziehungen

$$\delta_1(A) = d_1(A) = a_{11}\,, \quad \delta_n(A) = d_1(A) \cdot \ldots \cdot d_n(A) = \det(A)\,.$$

(5) *Die Matrix $C = LD^{\frac{1}{2}}$ heißt der* Cholesky-Faktor *von A. Es gilt*

$$A = LDL^t = (LD^{\frac{1}{2}}) \circ (D^{\frac{1}{2}}L^t) = CC^t\,.$$

Die Matrix $C \in \mathrm{L}_{\times}^{\llcorner}(n, \mathbb{R})$ ist eine positiv definite obere Dreiecksmatrix. Der Cholesky-Faktor C ist durch die Matrix A eindeutig bestimmt. Die Zerlegung $A = CC^t$ heißt die symmetrische Form der Cholesky-Zerlegung.

Beweis. Zu beweisen ist nur noch Aussage (4). Sei $1 \leq k < n$. Für beliebige Matrizen $A \in \mathbb{R}^{n\times n}$, $L \in \mathrm{L}^{\llcorner}(n, \mathbb{R})$ und $U \in \mathrm{U}^{\urcorner}(n, \mathbb{R})$ gelten

$$\begin{pmatrix} A_{1:k,1:k} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix} \circ A \circ \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix}\,,$$

$$\begin{pmatrix} L_{1:k,1:k} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix} \circ L\,,$$

$$\begin{pmatrix} U_{1:k,1:k} & 0 \\ 0 & 0 \end{pmatrix} = U \circ \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix}\,.$$

Aus diesen Beziehungen folgt (4). □

Beispiel 18.4. Es besteht die Cholesky-Zerlegung

$$A = \begin{pmatrix} 5 & 20 & -15 \\ 20 & 82 & -56 \\ -15 & -56 & 60 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{pmatrix} \begin{pmatrix} 1 & 4 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} .$$

Ohne vorher die Positivität der Hauptminoren geprüft zu haben, wenden wir das Verfahren zur Berechnung der strikten Bruhat-Zerlegung an. Wenn das Verfahren nicht vorzeitig abbricht, ist die Matrix A invertierbar. Wenn $P = E_3$ gilt und alle Diagonalelemente von D positiv sind, ist die gegebene quadratische reelle Matrix A tatächlich positiv definit. In diesem Fall gilt $U = L^t$.

$$\begin{array}{rrr|l}
5 & 20 & -15 & \\
20 & 82 & -56 & \\
-15 & -56 & 60 & \\
\hline
5 & 20 & -15 & Z_1 \\
0 & 2 & 4 & Z_2 - (+4)Z_1 \\
0 & 4 & 15 & Z_3 - (-3)Z_1 \\
\hline
5 & 20 & -15 & Z_1 \\
0 & 2 & 4 & Z_2 \\
0 & 0 & 7 & Z_3 - (+2)Z_2 \\
\hline\hline
1 & 4 & -3 & (+5)^{-1}Z_1 \\
0 & 1 & 2 & (+2)^{-1}Z_2 \\
0 & 0 & 1 & (+7)^{-1}Z_3 \\
\hline\hline
1 & 0 & 0 & \\
0 & 1 & 0 & \\
0 & 0 & 1 &
\end{array}$$

Aus der Sandwichform der Cholesky-Zerlegung lassen sich die Hauptminoren der positiv definiten Matrix A ablesen. Wir erhalten

$$\delta_1(A) = 5\,, \quad \delta_2(A) = 5 \cdot 2 = 10\,, \quad \delta_3(A) = 5 \cdot 2 \cdot 7 = 70\,.$$

Der Cholesky-Faktor von A ist die positiv definite untere Dreiecksmatrix

$$C = \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 4\sqrt{5} & \sqrt{2} & 0 \\ -3\sqrt{5} & 2\sqrt{2} & \sqrt{7} \end{pmatrix} .$$

□

19 Ergänzung: Orthogonale Spiegelungen

Satz und Definition 19.1. *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,\,.\,\rangle$ und der euklidischen Norm $\|\,.\,\|$ versehen. Sei $A \in \mathbb{R}^{n\times n}$ gegeben. Dann sind* (1) *bis* (3) *äquivalent.*

(1) $(\forall x \in \mathbb{R}^n): \ \|Ax\| = \|x\|$.

(2) $(\forall x, y \in \mathbb{R}^n): \ \langle Ax, Ay\rangle = \langle x, y\rangle$.

(3) $A^t A = E_n$.

Eine Matrix $A \in \mathbb{R}^{n\times n}$, die eine der äquivalenten Bedingungen (1) *bis* (3) *erfüllt heißt* orthogonal. *Sei* $\mathrm{O}(n,\mathbb{R})$ *die Menge der orthogonalen Matrizen $A \in \mathbb{R}^{n\times n}$. Es gelten* (4) *bis* (7).

(4) $E_n \in \mathrm{O}(n,\mathbb{R}) \subseteq \mathrm{GL}(n,\mathbb{R})$.

(5) $(\forall A, B \in \mathrm{O}(n,\mathbb{R})): \ A \circ B \in \mathrm{O}(n,\mathbb{R})$.

(6) $(\forall A \in \mathrm{O}(n,\mathbb{R})): \ A^{-1} = A^t \in \mathrm{O}(n,\mathbb{R})$.

(7) $(\forall A \in \mathrm{O}(n,\mathbb{R})): \ \det(A) = \pm 1$.

Nach (4), (5), (6) *ist* $\mathrm{O}(n,\mathbb{R})$ *eine Untergruppe von* $\mathrm{GL}(n,\mathbb{R})$.

Satz und Definition 19.2 (Orthogonale Spiegelungen)**.** *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,\,.\,\rangle$ und der euklidischen Norm $\|\,.\,\|$ versehen. Sei $w \in \mathbb{R}^n \setminus \{0\}$. Dann sei $S_w \in \mathbb{R}^{n\times n}$ die Matrix mit*

$$S_w = E_n - 2\,\frac{w \otimes w}{\|w\|^2}\,.$$

Es gelten (1) *bis* (8).

(1) *S_w ist symmetrisch mit der Spektralzerlegung*

$$S_w = \left(E_n - \frac{w \otimes w}{\|w\|^2}\right) - \frac{w \otimes w}{\|w\|^2}\,.$$

(2) $\sigma(S_w) = \{-1, 1\}$.

(3) $\ker(S_w + E_n) = [w]$.

(4) $\ker(S_w - E_3) = [w]^\perp$.

(5) $S_w^2 = S_w \circ S_w = E_n$.

(6) $S_w \in \mathrm{O}(n,\mathbb{R})$.

(7) $\det(S_w) = -1$.

(8) $\mathrm{tr}(S_w) = n - 2$.

Wegen dieser Eigenschaften heißt die Matrix $S_w \in \mathrm{O}(n,\mathbb{R})$ die orthogonale Spiegelung *an der Hyperebene*

$$H_w = [w]^\perp = \{x \in \mathbb{R}^n \mid \langle x, w\rangle = 0\}\,.$$

Eine orthogonale Matrix $A \in \mathrm{O}(n,\mathbb{R})$ heißt eine orthogonale Spiegelung*, wenn es einen normierten Vektor $v \in \mathbb{R}^n$ mit $A = S_v$ gibt.*

Beispiel 19.3. Die orthogonale Spiegelung an der ersten Winkelhalbierenden H der Ebene $\mathbb{R}^2$ vertauscht die beiden Koordinaten eines Punktes. Wir rechnen diesen elementargeometrischen Sachverhalt nach. Mit

$$H = H_w = \{\, x \in \mathbb{R}^2 \mid \langle x, w\rangle = 0\,\}\,, \quad w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \|w\|^2 = 2$$

erhalten wir

$$\begin{aligned} S_w = E_2 - 2\,\frac{w \otimes w}{\|w\|^2} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

Für alle $x = (x_1, x_2)^t \in \mathbb{R}^2$ gilt daher

$$S_w x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}.$$

Insbesondere folgt daraus $S_w^2 = S_w$. In Übereinstimmung mit 19.2 gelten

$$H = H_w = \ker(S_w - E_2)\,, \quad \det(S_w) = -1\,, \quad \mathrm{tr}(S_w) = 0\,.$$

□

Satz 19.4 (Élie Cartan, 1938)**.** *Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,\,.\,\rangle$ und der euklidischen Norm $\|\,.\,\|$ versehen. Jede Orthogonalmatrix $A \in \mathrm{O}(n,\mathbb{R})$ mit $A \neq E_n$ kann als Produkt von höchstens n orthogonalen Spiegelungen dargestellt werden.*

20 Drehungen. Euler-Winkel

Satz und Definition 20.1. *Sei* $n \in \mathbb{N}$.

(1) *Eine Basis* $\mathcal{V} = \{v_1, \dots, v_n\}$ *des* $\mathbb{R}^n$ *heißt* positiv orientiert, *wenn*

$$\det(v_1, \dots, v_n) = \det(V) > 0$$

gilt. Dabei ist $V = (v_1, \dots, v_n)$ *die Matrix mit den Spalten* $v_1, \dots, v_n$.

(2) *Eine Basis des* $\mathbb{R}^n$*, die nicht positiv orientiert ist, heißt* negativ orientiert.

(3) *Eine Matrix* $A \in \mathrm{GL}(n, \mathbb{R})$ *heißt* orientierungstreu, *wenn* $\det(A) > 0$ *gilt.*

(4) $A = (a_1, \dots, a_n) \in \mathrm{GL}(n, \mathbb{R})$ *ist genau dann orientierungstreu, wenn* $\mathcal{A} = \{a_1, \dots, a_n\}$ *eine positiv orientierte Basis des* $\mathbb{R}^n$ *ist.*

(5) *Die Menge der positiv orientierten Basen des* $\mathbb{R}^n$ *heißt die* Standard-Orientierung *des* $\mathbb{R}^n$. *Die Standard-Orientierung wird also mit Hilfe der Determinante und der Ordnungsstruktur der reellen Zahlen definiert.*

(6) *Die kanonische Basis* $\mathcal{E}_n = \{e_1, \dots, e_n\}$ *des* $\mathbb{R}^n$ *ist positiv orientiert und repräsentiert daher die Standard-Orientierung des* $\mathbb{R}^n$. *Wir sagen, dass der Vektorraum* $\mathbb{R}^n$ *mit der Standard-Orientierung versehen ist.*

Satz und Definition 20.2 (Drehmatrizen)**.** *Sei* $n \in \mathbb{N}$. *Der Vektorraum* $\mathbb{R}^n$ *sei mit dem euklidischen inneren Produkt* $\langle\, .\,,.\, \rangle$ *und der euklidischen Norm* $\|\,.\,\|$ *und der Standard-Orientierung versehen.*

(1) $A \in \mathrm{O}(n, \mathbb{R})$ *heißt eine* Drehmatrix, *wenn* $\det(A) = 1$ *gilt.*

(2) *Sei* $\mathrm{SO}(n, \mathbb{R})$ *die Menge der Drehmatrizen* $A \in \mathbb{R}^{n \times n}$.

(3) *Die Matrizen* $A \in \mathrm{SO}(2, \mathbb{R})$ *heißen* ebene Drehmatrizen.

(4) *Die Matrizen* $A \in \mathrm{SO}(3, \mathbb{R})$ *heißen* räumliche Drehmatrizen.

Es gelten die folgenden Ausagen (5) *bis* (7).

(5) $E_n \in \mathrm{SO}(n, \mathbb{R}) \subseteq \mathrm{O}(n, \mathbb{R}) \subseteq \mathrm{GL}(n, \mathbb{R})$.

(6) $(\forall A, B \in \mathrm{SO}(n, \mathbb{R}))$: $A \circ B \in \mathrm{SO}(n, \mathbb{R})$.

(7) $(\forall A \in \mathrm{SO}(n, \mathbb{R}))$: $A^{-1} = A^t \in \mathrm{SO}(n, \mathbb{R})$.

Nach (5), (6), (7) *ist* $\mathrm{SO}(n, \mathbb{R})$ *eine Untergruppe von* $\mathrm{O}(n, \mathbb{R})$ *und* $\mathrm{GL}(n, \mathbb{R})$.

Im Folgenden untersuchen wir ebene und räumliche Drehmatrizen.

Satz und Definition 20.3 (Ebene Drehmatrizen)**.** *Der Vektorraum* $\mathbb{R}^2$ *sei mit dem euklidischen inneren Produkt* $\langle\, .\,,.\, \rangle$ *und der euklidischen Norm* $\|\,.\,\|$ *und der Standard-Orientierung versehen. Dann gelten die folgenden Ausagen* (1) *bis* (3).

(1) $\mathrm{SO}(2,\mathbb{R})$ *besteht aus den reellen* (2×2)*-Matrizen der Form*

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

mit $a,\, b \in \mathbb{R}$ *und* $a^2 + b^2 = 1$*.*

(2) *Sei* $\measuredangle(e_1, Ae_1) \in [0,\pi]$ *der von den beiden Vektoren*

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^2\,, \quad Ae_1 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$$

eingeschlossene unorientierte Winkel. Siehe 3.11*. Dann gelten*

$$a = \cos(\measuredangle(e_1, Ae_1))\,, \quad |b| = \sin(\measuredangle(e_1, Ae_1))\,.$$

Der unorientierte Drehwinkel bestimmt die Diagonale von A. Wir heben hervor, dass der unorientierte Drehwinkel $\measuredangle(e_1, Ae_1)$ *nur die Beträge der Einträge außerhalb der Diagonale von A festlegt.*

(3) *Zu jedem* $A \in \mathrm{SO}(2,\mathbb{R})$ *gibt es ein eindeutig bestimmtes* $\varphi \in (-\pi,\pi]$ *mit*

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}\,.$$

Dabei heißt $\varphi \in (-\pi,\pi]$ *der* orientierte Drehwinkel *von A. Der orientierte Drehwinkel und ebene Drehmatrizen bestimmen einander wechselseitig.*

(4) *Für alle* $t \in \mathbb{R}$ *gilt*

$$D(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \in \mathrm{SO}(2,\mathbb{R})\,.$$

Offenbar gilt $D(t) = D(t+2\pi)$ *für alle* $t \in \mathbb{R}$*. Der reelle Parameter t heißt* Drehwinkel modulo 2π*.*

Wir behandeln nun die räumlichen Drehmatrizen. Um nicht jedesmal die Voraussetzungen wiederholen zu müssen, legen wir für die folgenden Sätze fest, dass der Vektorraum $\mathbb{R}^3$ mit dem euklidischen inneren Produkt $\langle\,.\,,.\,\rangle$ und der euklidischen Norm $\|\,.\,\|$ und der Standard-Orientierung versehen ist. Wir untersuchen zunächst Spektrum und Spur räumlicher Drehmatrizen. In Beweis von Satz 20.4 holen wir im Beweis von Satz 20.6 nach.

Satz 20.4. *Sei* $A \in \mathrm{SO}(3,\mathbb{R})$*. Dann gelten* (1) *bis* (7)*.*

(1) $\chi_A(\tau) = (-1)(\tau^3 - \mathrm{tr}(A)\,\tau^2 + \mathrm{tr}(A)\,\tau - 1) \in \mathbb{R}[\tau]\,.$

(2) $(\forall \lambda \in \sigma_{\mathbb{C}}(A)) :\; |\lambda| = 1\,.$

(3) $1 \in \sigma_{\mathbb{R}}(A)\,.$

(4) $A \neq E_3 \Leftrightarrow \dim(\ker(A - E_3)) = 1$.

(5) $\mathrm{tr}(A) \in [-1, 3]$.

(6) $\mathrm{tr}(A) \in (-1, 3) \Leftrightarrow A \neq A^t$.

(7) $\mathrm{tr}(A) = 3 \Leftrightarrow A = E_3$.

Beweis. Wir beweisen an dieser Stelle lediglich die Aussagen (1), (2), (3), (5). Die übrigen Aussagen (4), (6), (7) ergeben sich im Beweis des Satzes 20.6.

Sei $A \in \mathrm{SO}(3, \mathbb{R})$. Dann gelten

$$\det(A) = 1\,, \quad \mathrm{adj}(A) = A^{-1} = A^t\,.$$

Mit 14.10 erhalten wir

$$\begin{aligned}\chi_A(\tau) &= \det(A - \tau E_3) \\ &= (-1)(\tau^3 - \mathrm{tr}(A)\,\tau^2 + \mathrm{tr}(\mathrm{adj}(A))\,\tau - \det(A)) \\ &= (-1)(\tau^3 - \mathrm{tr}(A)\,\tau^2 + \mathrm{tr}(A)\,\tau - 1) \\ &= (-1)(\tau - 1)(\tau^2 - (1 - \mathrm{tr}(A))\,\tau + 1) \in \mathbb{R}[\tau]\,.\end{aligned}$$

Offenbar ist $\tau = 1$ eine Nullstelle von $\chi_A(\tau)$. Das Horner-Schema

	1	$-\mathrm{tr}(A)$	$\mathrm{tr}(A)$	-1
1		1	$1 - \mathrm{tr}(A)$	1
	1	$1 - \mathrm{tr}(A)$	1	0

liefert den quadratischen Faktor

$$q(\tau) = \tau^2 + (1 - \mathrm{tr}(A))\,\tau + 1 \in \mathbb{R}[\tau]\,.$$

Seien λ_1, λ_2, $\lambda_3 \in \mathbb{C}$ die Nullstellen von $\chi_A(\tau)$. Wir erhalten nacheinander die folgenden Beziehungen

$$\begin{aligned}\chi_A(\tau) &= (-1)(\tau - \lambda_1)(\tau - \lambda_2)(\tau - \lambda_3)\,, \quad \lambda_1 = 1\,, \\ \mathrm{tr}(A) &= \lambda_1 + \lambda_2 + \lambda_3\,, \quad \det(A) = \lambda_1\lambda_2\lambda_3 = 1\,. \\ q(\tau) &= (\tau - \lambda_2)(\tau - \lambda_3)\,, \quad \lambda_2\lambda_3 = 1\,, \\ \lambda_2 + \lambda_3 &= \mathrm{tr}(A) - \lambda_1 = \mathrm{tr}(A) - 1 \in \mathbb{R}\,, \\ \mathrm{Im}(\lambda_3) &= -\mathrm{Im}(\lambda_2)\,, \quad \mathrm{Re}(\lambda_3) = \mathrm{Re}(\lambda_2)\,, \quad |\lambda_2| = |\lambda_3| = 1\,.\end{aligned}$$

Daher gibt es genau ein $\psi \in (-\pi, \pi]$ mit

$$\lambda_2 = e^{i\psi}\,, \quad \lambda_3 = e^{-i\psi}\,.$$

Nun folgt

$$\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \lambda_3 = 1 + e^{i\psi} + e^{-i\psi} = 1 + 2\cos(\psi) \in [-1, 3]\,.$$

Wir weisen auf die Spurformel in Satz 20.6 hin. Dort wird der reelle Parameter ψ als Drehwinkel gedeutet. $\square$

Definition 20.5 (Drehachse, Drehebene)**.** *Sei* $A \in \mathrm{SO}(3,\mathbb{R})$.

(1) *Ein Teilvektorraum*

$$X \subseteq \ker(A - E_3) \subseteq \mathbb{R}^3$$

mit $\dim(X) = 1$ *heißt eine* Drehachse *von* A*. Der Orthogonalraum* $X^{\perp} \subseteq \mathbb{R}^3$ *heißt eine* Drehebene *von* A*.*

(2) *Im Fall* $A \neq E_3$ *ist der Eigenraum* $\ker(A - E_3)$ *eindimensional. Daher ist sind Drehachse und Drehebene eindeutig bestimmt.*

(3) *Im Fall* $A = E_3$ *ist jeder eindimensionale Teilraum des* $\mathbb{R}^3$ *eine Drehachse. Jeder zweidimensionale Teilraum ist dann eine Drehebene.*

Satz und Definition 20.6 (Euler. Räumliche Drehmatrizen. Drehformel, Drehvektor, Drehwinkel)**.**

(1) *Für* $v \in \mathbb{R}^3$ *mit* $\|v\| = 1$ *und* $\varphi \in \mathbb{R}$ *sei* $D(v,\phi) \in \mathbb{R}^{3\times 3}$ *die Matrix mit*

$$D(v,\varphi) = \left\{ \cos(\varphi)\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + (1-\cos(\varphi))\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix} \otimes \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix} \right.$$

$$\left. + \sin(\varphi)\begin{pmatrix}0&-v_3&v_2\\v_3&0&-v_1\\-v_2&v_1&0\end{pmatrix} \right\}.$$

Dann gelten:

(1.1) $D(v,\varphi) \in \mathrm{SO}(3,\mathbb{R})$.

(1.2) *Für alle* $x \in \mathbb{R}^3$ *gilt*

$$D(v,\varphi)\,x = \cos(\varphi)\,x + (1-\cos(\varphi))\langle x, v\rangle\, v + \sin(\varphi)\, v \times x\,.$$

(1.3) $\mathbb{R}v = [v]$ *ist eine Drehachse von* $D(v,\varphi)$.

(1.4) $\mathrm{tr}(D(v,\varphi)) = 1 + 2\cos(\varphi)$.

(1.5) $D(v,\varphi) = D(-v,-\varphi)$.

(1.6) $D(v,\varphi) = D(v,\varphi + 2\pi)$.

Der reelle Parameter φ *heißt der* Drehwinkel *modulo* 2π *der räumlichen Drehmatrix* $D(v,\varphi)$*.*

(2) *Sei* $A \in \mathrm{SO}(3,\mathbb{R})$.

(2.1) *Dann gibt es* $v \in \mathbb{R}^3$ *mit* $\|v\| = 1$ *und* $\varphi \in \mathbb{R}$ *mit*

$$A = D(v,\varphi)\,.$$

Diese Gleichung heißt Drehformel *oder* Normalform *einer räumlichen Drehmatrix.*

(2.2) *Es ist üblich, den Drehwinkel φ auf das Intervall $(-\pi, \pi]$ zu beschränken.*

(2.3) *Im Fall $A \neq A^t$ kann der normierte Vektor $v \in \ker(A - E_3)$ eindeutig durch die Bedingung festgelegt werden, dass $\varphi \in (0, \pi)$ gilt. Eine entsprechende Festlegung ist im Fall $A = A^t$ allerdings nicht möglich.*

(3) *Sei $A = A^t \in \mathrm{SO}(3, \mathbb{R})$.*

(3.1) *Sei $A = E_3$. Für alle $v \in \mathbb{R}^3$ mit $\|v\| = 1$ gilt*

$$A = D(v, 0) = E_3 .$$

Es gilt $\mathrm{tr}(A) = 3$. Die Spurformel (1.2) liefert

$$\cos(\varphi) = \tfrac{1}{2}(\mathrm{tr}(A) - 1) = 1 = \cos(0) .$$

(3.2) *Sei $A \neq E_3$. Für alle $v \in \ker(A - E_3)$ mit $\|v\| = 1$ gilt*

$$A = D(v, \pi) = D(-v, \pi) = -E_3 + 2v \otimes v = -S_v .$$

Dabei ist $S_v = S_{-v}$ die orthogonale Spiegelung an der Drehebene $[v]^\perp$. Siehe 19.2. Es gilt $\mathrm{tr}(A) = -1$. Die Spurformel (1.2) liefert

$$\cos(\varphi) = \tfrac{1}{2}(\mathrm{tr}(A) - 1) = -1 = \cos(\pi) .$$

(4) *Sei $A \in \mathrm{SO}(3, \mathbb{R})$ mit $A \neq A^t$. Dann gibt es ein eindeutig bestimmtes $v \in \mathbb{R}^3$ mit $\|v\| = 1$ und ein eindeutig bestimmtes $\varphi \in (0, \pi)$ derart, dass*

$$A = D(v, \varphi)$$

gilt. Dabei heißt v der normierte Drehvektor *von A. Der normierte Drehvektor legt die* Richtung *der Drehachse und damit den* Drehsinn *fest. Der Parameter $\varphi \in (0, \pi)$ heißt der* orientierte Drehwinkel *bezüglich v.*

(4.1) *Der orientierte Drehwinkel $\varphi \in (0, \pi)$ ist durch die* Spurformel

$$\mathrm{tr}(A) = 1 + 2\cos(\varphi)$$

eindeutig bestimmt.

(4.2) *Der normierte Drehvektor $v \in \ker(A - E_3)$ ergibt sich aus dem schiefsymmetrischen Anteil von A nach den Formeln*

$$\tfrac{1}{2}(A - A^t) = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} , \quad u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} , \quad v = \frac{1}{\|u\|}\, u .$$

Den Vektor u nennen wir den Drehvektor *von A.*

(4.3) *Der Sinus von $\varphi \in (0, \pi)$ ist durch*

$$\sin(\varphi) = \|u\| > 0$$

eindeutig bestimmt.

(4.4) (Rechte-Hand-Regel). *Sei $w \in [v]^\perp$ mit $\|w\| = 1$ beliebig gegeben. Dann ist $V = (v, w, v \times w) \in \mathrm{SO}(3, \mathbb{R})$ eine Drehmatrix mit*

$$V^{-1} \circ A \circ V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Außerdem ist $\{w, v \times w\}$ eine Orthonormalbasis der Drehebene $[v]^\perp$ der Drehmatrix A.

Beweis. Im ersten Schritt untersuchen wir räumliche Drehmatrizen der Form $D(v, \varphi)$. Im dritten Schritt zeigen wir, dass sich jede räumliche Drehmatrix in dieser Form schreiben lässt. Aus dieser Darstellung ergibt sich sofort ein Beweis des Satzes 20.4.

Erster Schritt. Seien $v = (v_1, v_2, v_3)^t \in \mathbb{R}^3$ mit $\|v\| = 1$ und $\varphi \in \mathbb{R}^3$ beliebig gegeben. Aus der Definition der Matrix $D(v, \varphi)$ folgen die Beziehungen

$$D(v, \varphi) = D(-v, -\varphi), \quad D(v, \varphi) = D(v, \varphi + 2\pi)$$

und die Spurformel

$$\mathrm{tr}(D(v, \varphi)) = 1 + 2\cos(\varphi).$$

Dabei verwenden wir, dass v ein normierter Vektor ist. Außerdem gilt

$$D(v, \varphi)\, x = \cos(\varphi)\, x + (1 - \cos(\varphi))\langle x, v\rangle\, v + \sin(\varphi)\, v \times x.$$

für alle $x \in \mathbb{R}^3$. Wir zeigen, dass $D(v, \varphi) \in \mathbb{R}^{3\times 3}$ eine Drehmatrix ist. Es gilt

$$v \times v = 0_3.$$

Folglich ist v ein normierter Eigenvektor von $D(v, \varphi)$ zum Eigenwert 1. Wir wählen irgendeinen Vektor $w \in \mathbb{R}^3 \setminus \{0\}$ mit $\|w\| = 1$. Dann ist

$$\mathcal{V} = \{v, w, v \times w\} \subseteq \mathbb{R}^3$$

eine positiv orientierte Orthonormalbasis des $\mathbb{R}^3$. Es gilt

$$v \times (v \times w) = -w.$$

Spaltenweises Auswerten ergibt, dass

$$V = (v, w, v \times w) \in \mathrm{SO}(3, \mathbb{R})$$

die übersichtliche Beziehung

$$D(v, \varphi) \circ V = V \circ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

liefert. Aus dieser Beziehung folgt mit Satz 14.6 die Behauptung

$$D(v, \varphi) \in \mathrm{SO}(3, \mathbb{R})\,.$$

Damit sind alle Aussagen des Abschnittes (1) bewiesen.

Zweiter Schritt. Für $v = e_1$ und $V = (e_1, e_2, e_3) = E_3$ ergibt sich

$$D(e_1, \varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}\,.$$

Damit gilt

$$D(v, \varphi) = V \circ D(e_1, \varphi) \circ V^t\,.$$

Wir weisen auf 20.9 und Satz 20.10 hin.

Dritter Schritt. Im Folgenden sei $A \in \mathrm{SO}(3, \mathbb{R})$. Dann gilt

$$\begin{aligned} \chi_A(\tau) &= (-1)(\tau^3 - \mathrm{tr}(A)\,\tau^2 + \mathrm{tr}(A)\,\tau - 1) \\ &= (-1)(\tau - 1)(\tau^2 + (1 - \mathrm{tr}(A))\,\tau + 1)\,. \end{aligned}$$

Siehe Satz 14.7 und Beispiel 14.10. Also gibt es $v \in \mathbb{R}^3 \setminus \{0\}$ mit

$$Av = v\,, \quad \|v\| = 1\,.$$

Wir wählen $w \in [v]^\perp$ mit $\|w\| = 1$. Dann ist $V = (v, w, v \times w) \in \mathrm{SO}(3, \mathbb{R})$ eine Drehmatrix mit

$$V^t \circ A \circ V = \begin{pmatrix} 1 & 0_{1\times 2} \\ 0_{2\times 1} & M \end{pmatrix}\,,$$

wobei $M \in \mathrm{SO}(2, \mathbb{R})$ gilt. Nach Satz 20.3 gibt einen eindeutig bestimmten reellen Parameter $\varphi \in (-\pi, \pi]$ mit

$$M = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}\,.$$

Mit den Überlegungen im ersten Schritt erhalten wir

$$A = V \circ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix} \circ V^t = D(v, \varphi)\,.$$

Wir ziehen einige Folgerungen aus dieser Beziehung.

Vierter Schritt. Mit Satz 14.6 folgt die Spurformel

$$\operatorname{tr}(A) = 1 + 2\cos(\varphi)\,.$$

Offenbar ist A genau dann symmetrisch, wenn $\sin(\varphi) = 0$ gilt. In diesem Fall gelten

$$A = V \circ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \circ V^t = E_3$$

oder

$$\begin{aligned} A &= V \circ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \circ V^t \\ &= -E_3 + V \circ \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \circ V^t \\ &= -E_3 + 2\, v \otimes v = -S_v\,. \end{aligned}$$

Also sind

$$E_3\,, \quad -E_3 + 2\, v \otimes v\,,$$

wobei $v \in \mathbb{R}^3 \setminus \{0\}$ ein normierter Vektor ist, die einzigen symmetrischen Matrizen in $\mathrm{SO}(3,\mathbb{R})$. Beide Typen symmetrischer Drehmatrizen lassen sich durch ihre Spur unterschieden.

Fünfter Schritt. Nun folgen alle Aussagen des Satzes 20.4 und die Aussagen in den Abschnitten (2), (3), (4) des Satzes 20.6. $\square$

Beispiel 20.7. Die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 \\ \frac{4}{5} & -\frac{3}{5} & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

eine symmetrische Drehmatrix. Nach Satz 20.6 ist $\lambda_1 = 1$ ein einfacher Eigenwert und $\lambda_2 = \lambda_3 = -1$ ein doppelter Eigenwert von A. Die Vektoren

$$v_1 = \tfrac{1}{5}\sqrt{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}\,, \quad v_2 = \tfrac{1}{5}\sqrt{5} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}\,, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

bilden eine Orthogonalbasis aus Eigenvektoren von A. Dabei gilt $Av_k = \lambda_k v_k$ für $k = 1, 2, 3$. Die Matrix A beschreibt eine demnach Drehung um die Drehachse $[v_1]$ mit dem Drehwinkel π. Wir rechnen nach, dass $-A$ die orthogonale Spiegelung an der Drehebene $[v_1]^\perp = \ker(A - E_3)$ ist. Es gilt

$$
\begin{aligned}
S_{v_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \tfrac{2}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \\
&= \tfrac{1}{5} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} - \tfrac{2}{5} \begin{pmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&= \tfrac{1}{5} \begin{pmatrix} -3 & -4 & 0 \\ -4 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} -\frac{3}{5} & -\frac{4}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} = -A\,.
\end{aligned}
$$

Offenbar gilt $A = S_{v_1} = S_{-v_1}$. $\square$

Beispiel 20.8. Die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$
A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}
$$

ist eine Drehmatrix. Es gilt $A \neq A^t$. Wir berechnen den normierten Drehvektor $v \in \mathbb{R}^3$ sowie Cosinus und Sinus des orientierten Drehwinkel $\varphi \in (0, \pi)$. Es gilt

$$
\tfrac{1}{2}(A - A^t) = \begin{pmatrix} 0 & 0 & \frac{2}{3} \\ 0 & 0 & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}.
$$

Damit erhalten wir

$$
u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ 0 \end{pmatrix}, \quad \|u\|^2 = \tfrac{8}{9}, \quad v = \frac{1}{\|u\|}\, u = \tfrac{1}{2}\sqrt{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.
$$

Zur Berechnung von $\cos(\varphi)$ werten wir die Spurformel aus. Wir erhalten

$$
\operatorname{tr}(A) = \tfrac{1}{3}, \quad \cos(\varphi) = -\tfrac{1}{3}, \quad \sin(\varphi) = \tfrac{2}{3}\sqrt{2} = \|u\|\,.
$$

$\square$

Beispiele 20.9. Sei $\varphi \in \mathbb{R}$. Für die kanonischen Basisvektoren e_1, e_2, $e_3 \in \mathbb{R}^3$ erhalten wir die Drehmatrizen

$$D(e_1, \varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix},$$

$$D(e_2, \varphi) = \begin{pmatrix} \cos(\varphi) & 0 & -\sin(\varphi) \\ 0 & 1 & 0 \\ \sin(\varphi) & 0 & \cos(\varphi) \end{pmatrix},$$

$$D(e_3, \varphi) = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

□

Satz und Definition 20.10 (Euler-Zerlegung, Euler-Winkel)**.**

(1) *Jede Drehmatrix* $A \in \mathrm{SO}(3, \mathbb{R})$ *besitzt eine Produktdarstellung*

$$A = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi)$$

mit ϕ, θ, $\psi \in \mathbb{R}$. *Eine solche Produktdarstellung heißt eine* Euler-Zerlegung *von* A. *Die reellen Parameter* ϕ, θ, ψ *heißen* Euler-Winkel *von* A.

(2) *Sei* $A = (a_{\nu\mu})_{\nu,\mu=1,2,3} \in \mathrm{SO}(3, \mathbb{R})$ *eine Drehmatrix mit den Euler-Winkeln* ψ, θ, $\phi \in \mathbb{R}$. *Dann gilt*

$$A = \begin{pmatrix} a_{11} & a_{12} & \sin(\phi)\sin(\theta) \\ a_{21} & a_{22} & -\cos(\phi)\sin(\theta) \\ \sin(\theta)\sin(\psi) & \sin(\theta)\cos(\psi) & \cos(\theta) \end{pmatrix}$$

mit

$$\begin{aligned} a_{11} &= \cos(\phi)\cos(\psi) - \sin(\phi)\cos(\theta)\sin(\psi)\,, \\ a_{12} &= -\cos(\phi)\sin(\psi) - \sin(\phi)\cos(\theta)\cos(\psi)\,, \\ a_{21} &= \sin(\phi)\cos(\psi) + \cos(\phi)\cos(\theta)\sin(\psi)\,, \\ a_{22} &= -\sin(\phi)\sin(\psi) + \cos(\phi)\cos(\theta)\cos(\psi)\,. \end{aligned}$$

Daher können die Einträge a_{11}, a_{12}, a_{21}, a_{22} *aus den den übrigen Einträgen der Drehmatrix* A *berechnet werden. Siehe* 20.11.

(3) *Im Fall* $Ae_3 \neq \pm e_3$ *sind die Euler-Winkel durch die Bedingungen*

$$\phi \in [0, 2\pi)\,, \quad \theta \in (0, \pi)\,, \quad \psi \in [0, 2\pi)$$

eindeutig bestimmt.

Siehe die folgenden Sätze 20.12 *und* 20.13 *zur Berechnung und geometrischen Deutung der Euler-Winkel.*

Schema 20.11.

$$D(e_3,\phi)\circ D(e_1,\theta)\circ D(e_3,\psi) = \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \cos(\theta)\sin(\psi) & \cos(\theta)\cos(\psi) & -\sin(\theta) \\ \sin(\theta)\sin(\psi) & \sin(\theta)\cos(\psi) & \cos(\theta \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\phi)\cos(\psi)-\sin(\phi)\cos(\theta)\sin(\psi) & -\cos(\phi)\sin(\psi)-\sin(\phi)\cos(\theta)\cos(\psi) & \sin(\phi)\sin(\theta) \\ \sin(\phi)\cos(\psi)+\cos(\phi)\cos(\theta)\sin(\psi) & -\sin(\phi)\sin(\psi)+\cos(\phi)\cos(\theta)\cos(\psi) & -\cos(\phi)\sin(\theta) \\ \sin(\theta)\sin(\psi) & \sin(\theta)\cos(\psi) & \cos(\theta) \end{pmatrix}.$$

Satz 20.12 (Berechnung der Euler-Winkel)**.** *Sei $A = (a_{\nu\mu}) \in \mathrm{SO}(3, \mathbb{R})$ gegeben. Im Übrigen verwenden wir die Bezeichnungen aus* 20.10.

(1) *Stets ist $\theta \in [0, \pi]$ durch die Gleichung*

$$\cos(\theta) = a_{33}$$

eindeutig bestimmt. Dann gilt

$$\sin(\theta) = \sqrt{1 - a_{33}^2} \geq 0\,.$$

(2) *Im Fall $a_{33} \neq \pm 1$ ist $\phi \in [0, 2\pi)$ ist durch die beiden Gleichungen*

$$\cos(\phi) = -\frac{a_{23}}{\sqrt{1 - a_{33}^2}}\,, \quad \sin(\phi) = \frac{a_{13}}{\sqrt{1 - a_{33}^2}}$$

eindeutig bestimmt.

(3) *Im Fall $a_{33} \neq \pm 1$ ist $\psi \in [0, 2\pi)$ ist durch die beiden Gleichungen*

$$\cos(\psi) = \frac{a_{32}}{\sqrt{1 - a_{33}^2}}\,, \quad \sin(\psi) = \frac{a_{31}}{\sqrt{1 - a_{33}^2}}$$

eindeutig bestimmt.

(4) *Im Fall $a_{33} = 1$ gibt es a, $b \in \mathbb{R}$ mit $a^2 + b^2 = 1$ und*

$$A = \begin{pmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & 1 \end{pmatrix}\,, \quad a = \cos(\phi + \psi)\,, \quad b = \sin(\phi + \psi)\,.$$

Die Matrix A beschreibt eine Drehung um die Drehachse $\mathbb{R}e_3$ mit dem Drehwinkel $\varphi = \phi + \psi$. Es besteht die Zerlegung

$$A = D(e_3, \phi) \circ D(e_1, 0) \circ D(e_3, \psi) = D(e_3, \phi + \psi)\,.$$

Die linke obere (2×2)-Teilmatrix von A ist eine ebene Drehmatrix.

(5) *Im Fall $a_{33} = -1$ ist A symmetrisch. Es gibt a, $b \in \mathbb{R}$ mit $a^2 + b^2 = 1$ und*

$$A = \begin{pmatrix} a & b & 0 \\ b & -a & 0 \\ 0 & 0 & -1 \end{pmatrix}\,, \quad a = \cos(\phi - \psi)\,, \quad b = \sin(\phi - \psi)\,.$$

Die Matrix A beschreibt eine Drehung mit dem Drehwinkel $\varphi = \pi$. Siehe Satz 20.6*, Aussage* (3.2)*. Die obere linke (2×2)-Teilmatrix von A ist keine ebene Drehmatrix. Es besteht die Zerlegung*

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ -\sin(\psi) & -\cos(\psi) \end{pmatrix}$$

in orthogonale Matrizen. Dabei ist die (2×2)-Matrix, die den Euler-Winkel ϕ enthält, eine ebene Drehmatrix.

Satz und Definition 20.13 (Geometrische Deutung der Euler-Winkel)**.** *Sei*

$$A = (a_1, a_2, a_3) \in \mathrm{SO}(3, \mathbb{R})$$

mit $a_3 \neq \pm e_3$ *gegeben. Sei* $\mathcal{A} = \{a_1, a_2, a_3\}$ *die positiv orientierte Orthonormalbasis des* $\mathbb{R}^3$*, die aus den Spalten von* A *besteht.*

(1) *Der Einheitsvektor* $e_N \in \mathbb{R}^3$ *mit*

$$e_N = \frac{e_3 \times a_3}{\|e_3 \times a_3\|}$$

heißt der Knotenvektor *von* A*. Es gilt*

$$0 \in \mathbb{R}e_N = H_{e_3,0} \cap H_{a_3,0} \subseteq \mathbb{R}^3 .$$

Die Gerade $\mathbb{R}e_N$ *heißt die* Knotenlinie *von* A*.*

(2) *Für alle*

$$\phi \in [0, 2\pi) , \quad \theta \in (0, \pi) , \quad \psi \in [0, 2\pi)$$

sind die geometrischen Bedingungen

$$D(e_3, \phi)e_1 = e_N , \quad D(e_N, \theta)e_3 = a_3 , \quad D(a_3, \psi)e_N = a_1 \tag{T}$$

genau dann erfüllt, wenn die Euler-Zerlegung

$$A = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi) . \tag{E1}$$

besteht. In diesem Fall gilt

$$A = D(a_3, \psi) \circ D(e_N, \theta) \circ D(e_3, \phi) . \tag{E2}$$

Wir heben die unterschiedliche Position der drei Parameter ϕ*,* θ*,* ψ *in den Zerlegungen* (E1) *und* (E2) *hervor. Siehe* (E3) *in Aussage* (6)*.*

(3) *Die Euler-Winkel*

$$\phi \in [0, 2\pi) , \quad \theta \in (0, \pi) , \quad \psi \in [0, 2\pi)$$

von A *lassen sich aus den Gleichungen*

$$\begin{aligned} \cos(\phi) &= \langle e_N, e_1 \rangle , & \sin(\phi) &= \langle e_N, e_2 \rangle , \\ \cos(\theta) &= \langle e_3, a_3 \rangle , & \sin(\theta) &= \|e_3 \times a_3\| , \\ \cos(\psi) &= \langle e_N, a_1 \rangle , & \sin(\psi) &= -\langle e_N, a_2 \rangle \end{aligned}$$

berechnen.

(4) *Die Ebenen* $H_{e_3,0}$ *und* $H_{a_3,0}$ *schneiden sich längs der Knotenlinie* $\mathbb{R}e_N$ *unter dem Winkel* $\theta = \measuredangle(e_3, a_3) \in (0, \pi)$*. Es gelten*

$$\cos(\theta) = \langle e_3, a_3 \rangle \in (-1, 1) , \quad \sin(\theta) = \|e_3 \times a_3\| \in (0, 1) .$$

(5) *Die Euler-Winkel $\phi \in [0, 2\pi)$ und $\psi \in [0, 2\pi)$ erfüllen*

$$e_N = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix}, \quad (e_N)_{\mathcal{A}} = A^{-1} e_N = \begin{pmatrix} \cos(\psi) \\ -\sin(\psi) \\ 0 \end{pmatrix}.$$

Dabei ist $(e_N)_{\mathcal{A}}$ der Koordinatenvektor von e_N bezüglich $\mathcal{A}$.

(6) *Für alle $x \in \mathbb{R}^3$ gilt*

$$x_{\mathcal{A}} = A^{-1}x = D(e_3, -\psi) \circ D(e_1, -\theta) \circ D(e_3, -\phi)\, x. \tag{E3}$$

Dabei ist $x_{\mathcal{A}}$ der Koordinatenvektor von x bezüglich $\mathcal{A}$.

Beweis. Sei $A = (a_1, a_2, a_3) \in \mathrm{SO}(3, \mathbb{R})$ mit $a_3 \neq \pm e_3$ gegeben. Wegen $a_3 \neq \pm e_3$ gilt $e_3 \times a_3 \neq 0$. Also ist e_N ein Einheitsvektor mit

$$e_N \perp e_3, \quad e_N \perp a_3.$$

Die Ebenen $H_{e_3,0}$ und $H_{a_3,0}$ schneiden sich längs der Knotenlinie $\mathbb{R}e_N$ unter dem Winkel $\measuredangle(e_3, a_3) \in (0, \pi)$. Wir betrachten die folgenden vier positiv orientierten Orthonormalbasen

$$\begin{aligned} \mathcal{E} &= \{e_1, e_2, e_3\}, \\ \mathcal{V} &= \{e_N, e_3 \times e_N, e_3\}, \\ \mathcal{W} &= \{e_N, a_3 \times e_N, a_3\}, \\ \mathcal{A} &= \{a_1, a_2, a_3\}. \end{aligned}$$

Erster Schritt. Es seien $\phi \in [0, 2\pi)$, $\theta \in (0, \pi)$, $\psi \in [0, 2\pi)$ mit

$$D(e_3, \phi)e_1 = e_N, \quad D(e_N, \theta)e_3 = a_3, \quad D(a_3, \psi)e_N = a_1 \tag{T}$$

gegeben. Wir zeigen, dass ϕ, θ, ψ die Euler-Winkel von A sind. Aus

$$\begin{aligned} D(e_3, \phi)&\colon \{e_1, e_2, e_3\} \mapsto \{e_N, e_3 \times e_N, e_3\}, \\ D(e_N, \theta)&\colon \{e_N, e_3 \times e_N, e_3\} \mapsto \{e_N, a_3 \times e_N, a_3\}, \\ D(a_3, \psi)&\colon \{e_N, a_3 \times e_N, a_3\} \mapsto \{a_1, a_2, a_3\} \end{aligned}$$

folgt die multiplikative Zerlegung

$$A = D(a_3, \psi) \circ D(e_N, \theta) \circ D(e_3, \phi).$$

Außerdem ergeben sich die Beziehungen

$$\begin{aligned} x_{\mathcal{V}} &= M_{\mathcal{E}\mathcal{V}}\, x_{\mathcal{E}} = D(e_3, \phi)^{-1}x, \\ x_{\mathcal{W}} &= M_{\mathcal{V}\mathcal{W}}\, x_{\mathcal{V}} = D(e_1, \theta)^{-1}x_{\mathcal{V}} = D(e_1, \theta)^{-1}D(e_3, \phi)^{-1}x, \\ x_{\mathcal{A}} &= M_{\mathcal{W}\mathcal{A}}\, x_{\mathcal{W}} = D(e_3, \psi)^{-1}x_{\mathcal{W}} = D(e_3, \psi)^{-1}D(e_1, \theta)^{-1}D(e_3, \phi)^{-1}x, \\ x_{\mathcal{A}} &= M_{\mathcal{E}\mathcal{A}}\, x_{\mathcal{E}} = A^{-1}x \end{aligned}$$

für alle $x \in \mathbb{R}^3$. Folglich gilt die Euler-Zerlegung

$$A = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi)\,. \tag{E}$$

Aus den geometrischen Bedingungen (T) folgt demnach, dass die rellen Parameter ϕ, θ, ψ die Euler-Winkel im Sinne von Satz und Definition 20.10 sind.

Zweiter Schritt. Es seien $\phi \in [0, 2\pi)$, $\theta \in (0, \pi)$, $\psi \in [0, 2\pi)$ derart gegeben, dass die Euler-Zerlegung (E) gilt. Wir zeigen, dass ϕ, θ, ψ die geometrischen Bedingungen (T) erfüllen. Dabei verwenden wir mehrfach die Darstellung aus 20.10 der Einträge $a_{\nu\mu}$ der Drehmatrix A als Funktionen von ϕ, θ, ψ.

Zur Vorbereitung berechnen wir die Koordinatenvektoren des Knotenvektors e_N bezüglich der Orthonormalbasen $\mathcal{E}$ und $\mathcal{A}$ in Abhängigkeit der Euler-Winkel ϕ und ψ. Nach 20.10 gilt

$$e_3 \times a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} \sin(\phi)\sin(\theta) \\ -\cos(\phi)\sin(\theta) \\ \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\phi)\sin(\theta) \\ \sin(\phi)\sin(\theta) \\ 0 \end{pmatrix}.$$

Normieren ergibt

$$(e_N)_{\mathcal{E}} = e_N = \frac{e_3 \times a_3}{\|e_3 \times a_3\|} = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix}.$$

Außerdem gelten

$$\langle e_3, a_3 \rangle = \cos(\theta)\,, \quad \|e_3 \times a_3\| = \sin(\theta) > 0\,.$$

Insbesondere erhalten wir

$$\theta = \measuredangle(e_3, a_3) \in (0, \pi)\,.$$

Aus 20.10 folgen die Beziehungen

$$\langle e_N, a_1 \rangle = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \cos(\phi)\cos(\psi) - \sin(\phi)\cos(\theta)\sin(\psi) \\ \sin(\phi)\cos(\psi) + \cos(\phi)\cos(\theta)\sin(\psi) \\ \sin(\theta)\sin(\psi) \end{pmatrix} = \cos(\psi)\,,$$

$$\langle e_N, a_2 \rangle = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -\cos(\phi)\sin(\psi) - \sin(\phi)\cos(\theta)\cos(\psi) \\ -\sin(\phi)\sin(\psi) + \cos(\phi)\cos(\theta)\cos(\psi) \\ \sin(\theta)\cos(\psi) \end{pmatrix} = -\sin(\psi)\,,$$

$$\langle e_N, a_3 \rangle = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \sin(\phi)\sin(\theta) \\ -\cos(\phi)\sin(\theta) \\ \cos(\theta) \end{pmatrix} = 0\,.$$

Diese Beziehungen lassen sich in der Form

$$(e_N)_{\mathcal{A}} = A^{-1} e_N = \begin{pmatrix} \cos(\psi) \\ -\sin(\psi) \\ 0 \end{pmatrix}$$

zusammenfassen.

Wir wenden uns nun dem Nachweis von (T) zu. Erstens untersuchen wir die Wirkung der Drehmatrix $D(e_3, \phi)$ auf die Basis $\mathcal{E} = \{e_1, e_2, e_3\}$. Wegen

$$e_3 = D(e_3,\phi)e_3\,, \quad e_N = \begin{pmatrix}\cos(\phi)\\ \sin(\phi)\\ 0\end{pmatrix} = D(e_3,\phi)e_1\,,$$

$$e_3 \times e_N = \begin{pmatrix}0\\0\\1\end{pmatrix} \times \begin{pmatrix}\cos(\phi)\\ \sin(\phi)\\ 0\end{pmatrix} = \begin{pmatrix}-\sin(\phi)\\ \cos(\phi)\\ 0\end{pmatrix} = D(e_3,\phi)e_2$$

gilt

$$D(e_3,\phi)\colon \{e_1, e_2, e_3\} \mapsto \{e_N, e_3 \times e_N, e_3\}\,.$$

Zweitens untersuchen wir die Wirkung von $D(e_N, \theta)$ auf die positiv orientierte Orthonormalbasis $\mathcal{V} = \{e_N, e_3 \times e_N, e_3\}$. Wegen

$$D(e_N,\theta)e_3 = \cos(\theta)\, e_3 + \sin(\theta)\, e_N \times e_3 = \begin{pmatrix}\sin(\phi)\sin(\theta)\\ -\cos(\phi)\sin(\theta)\\ \cos(\theta)\end{pmatrix} = a_3\,,$$

$$D(e_N,\theta)e_N = e_N = \begin{pmatrix}\cos(\phi)\\ \sin(\phi)\\ 0\end{pmatrix}\,,$$

$$D(e_N,\theta)(e_3 \times e_N) = a_3 \times e_N = \begin{pmatrix}-\sin(\phi)\cos(\theta)\\ \cos(\phi)\cos(\theta)\\ \sin(\theta)\end{pmatrix}$$

gilt

$$D(e_N,\theta)\colon \{e_N, e_3 \times e_N, e_3\} \mapsto \{e_N, a_3 \times e_N, a_3\}\,.$$

Drittens untersuchen wir die Wirkung von $D(a_3, \psi)$ auf die positiv orientierte Orthonormalbasis $\{e_N, a_3 \times e_N, a_3\}$. Wegen

$$\begin{aligned} D(a_3,\psi)e_N &= \cos(\psi)\, e_N + \sin(\psi)\, a_3 \times e_N \\ &= \begin{pmatrix}\cos(\phi)\cos(\psi)\\ \sin(\phi)\cos(\psi)\\ 0\end{pmatrix} + \begin{pmatrix}-\sin(\phi)\cos(\theta)\sin(\psi)\\ \cos(\phi)\cos(\theta)\sin(\psi)\\ \sin(\theta)\sin(\psi)\end{pmatrix} = a_1\,, \\ D(a_3,\psi)a_3 &= a_3\,, \quad D(a_3,\psi)(a_3 \times e_N) = a_3 \times a_1 = a_2 \end{aligned}$$

gilt

$$D(a_3,\psi)\colon \{e_N, a_3 \times e_N, a_3\} \mapsto \{a_1, a_2, a_3\}\,.$$

Also gelten die geometrischen Bedingungen (T). Damit ist der Beweis von 20.13 abgeschlossen. □

Schema 20.14.

$$A = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi) \in \mathrm{SO}(3, \mathbb{R}), \quad Ae_3 \neq \pm e_3 .$$

$$\phi \in [0, 2\pi), \quad \psi \in [0, 2\pi).$$

$$A \begin{pmatrix} \cos(\psi) \\ -\sin(\psi) \\ 0 \end{pmatrix} \quad = \quad \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix} \quad = \quad \frac{e_3 \times Ae_3}{\|e_3 \times Ae_3\|} \quad = \quad \begin{array}{c} e_1 \\ \downarrow \; D(e_3, \phi) \\ e_N \\ \downarrow \; D(Ae_3, \psi) \\ Ae_1 \end{array}$$

$$e_N = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix}, \quad (e_N)_{\mathcal{A}} = \begin{pmatrix} \cos(\psi) \\ -\sin(\psi) \\ 0 \end{pmatrix}.$$

$$\theta = \measuredangle(e_3, Ae_3) \in (0, \pi).$$

$$\begin{array}{c} e_3 \\ \downarrow \; D(e_N, \theta) \\ Ae_3 \end{array}$$

Beispiel 20.15. Sei $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

die Drehmatrix aus Beispiel 20.8. Offenbar gilt $Ae_3 \neq \pm e_3$. Also sind die Euler-Winkel ϕ, θ, ψ durch die Bedingungen

$$\phi \in [0, 2\pi)\,, \quad \theta \in (0, \pi)\,, \quad \psi \in [0, 2\pi)$$

eindeutig bestimmt.

Erste Methode. Mit den Formeln aus 20.12 erhalten wir

$$\begin{aligned} \cos(\theta) &= -\tfrac{1}{3}\,, & \sin(\theta) &= \tfrac{2}{3}\sqrt{2}\,, \\ \cos(\phi) &= \tfrac{1}{2}\sqrt{2}\,, & \sin(\phi) &= \tfrac{1}{2}\sqrt{2}\,, \\ \cos(\psi) &= \tfrac{1}{2}\sqrt{2}\,, & \sin(\psi) &= -\tfrac{1}{2}\sqrt{2}\,. \end{aligned}$$

Zweite Methode. Wir wenden Formeln aus 20.13 an. Zuerst berechnen wir $\cos(\theta)$ als inneres Produkt von e_3 und Ae_3. Dann berechnen wir den Knotenvektor e_N. Dabei erhalten wir $\sin(\theta)$ als euklidische Norm von $e_3 \times Ae_3$. Die ersten beiden Komponenten von e_N sind $\cos(\phi)$ und $\sin(\phi)$.

$$\cos(\theta) = \langle e_3, Ae_3 \rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix} = -\tfrac{1}{3}\,,$$

$$e_3 \times Ae_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ 0 \end{pmatrix}\,,$$

$$\|e_3 \times Ae_3\| = \tfrac{2}{3}\sqrt{2} = \sin(\theta)\,,$$

$$\begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix} = e_N = \frac{e_3 \times Ae_3}{\|e_3 \times Ae_3\|} = \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix}\,.$$

Wir berechnen $\cos(\psi)$ und $-\sin(\psi)$ als die erste und zweite Komponente des Vektors $A^t e_N$.

$$\begin{pmatrix} \cos(\psi) \\ -\sin(\psi) \\ 0 \end{pmatrix} = A^t e_N = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix}\,.$$

Probe. Zum Abschluss rechnen wir nach, dass im vorliegenden Fall tatsächlich

$$A = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi)$$

gilt. Einsetzen und Ausmultiplizieren ergibt

$$
\begin{aligned}
&D(e_3,\phi)\circ D(e_1,\theta)\circ D(e_3,\psi)\\
&=\begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0\\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0\\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0\\ 0 & -\frac{1}{3} & -\frac{2}{3}\sqrt{2}\\ 0 & \frac{2}{3}\sqrt{2} & -\frac{1}{3} \end{pmatrix}
\begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0\\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0\\ 0 & 0 & 1 \end{pmatrix}\\
&=\begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0\\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0\\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0\\ \frac{1}{6}\sqrt{2} & -\frac{1}{6}\sqrt{2} & -\frac{2}{3}\sqrt{2}\\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}\\
&=\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix} = A\,.
\end{aligned}
$$

□

21 Ergänzung: Gruppen

Definition 21.1. *Ein Paar* $(G, \diamond)$ *bestehend aus einer nicht-leeren Menge* G *und einer Abbildung* $\diamond : G \times G \to G$, $(g, h) \mapsto g \diamond h$ *heißt eine* Gruppe, *wenn die folgenden Bedingungen* (1) *und* (2) *erfüllt sind.*

(1) *Es gilt das* Assoziativgesetz

$$(\forall f, g, h \in G): \quad f \diamond (g \diamond h) = (f \diamond g) \diamond h\,.$$

(2) *Es gibt ein eindeutig bestimmtes Element* $\eta \in G$ *mit den Eigenschaften* (i) *und* (ii).

   (i) *Das Element* η *ist ein* neutrales Element *bezüglich der* Gruppenoperation $\diamond$, *das heißt, es gilt*

$$(\forall g \in G): \quad \eta \diamond g = g \diamond \eta = g\,.$$

   (ii) *Zu jedem* $g \in G$ *gibt es ein eindeutig bestimmtes* inverses Element $g' \in G$ *bezüglich* η *und* $\diamond$, *das heißt, es gilt*

$$(\forall g \in G)(\exists g' \in G): \quad g' \diamond g = g \diamond g' = \eta\,.$$

Das Element η *heißt das* neutrale Element *der Gruppe* $(G, \diamond)$.

Wenn außerdem die Bedingung (3) *erfüllt ist, dann heißt das Paar* $(G, \diamond)$ *eine* kommutative *oder* abelsche *Gruppe. Andernfalls heißt die Gruppe* nicht-abelsch *oder* nicht-kommutativ.

(3) *Es gilt das* Kommutativgesetz

$$(\forall f, g \in G): \quad f \diamond g = g \diamond f\,.$$

Wenn keine Verwechslungen zu befürchten sind, schreiben wir gh *anstelle von* $g \diamond h$ *und sagen wir vereinfachend, dass* G *eine Gruppe ist.*

Manchmal wird die Gruppenoperation mit einem Multiplikationspunkt $\cdot$ oder dem Kompositionszeichen $\circ$ bezeichnet. In diesem Fall sprechen wir von einer *multiplikativ geschriebenen* Gruppe. Das neutrale Element einer multiplikativ geschriebenen Gruppe wird *Eins* genannt und mit 1 oder e oder E bezeichnet. Das inverse Element $g' \in G$ wird dann mit g^{-1} bezeichnet.

Die Gruppenoperation einer kommutativen Gruppe wird oft mit einem Pluszeichen $+$ bezeichnet. In diesem Fall sprechen wir von einer *additiv geschriebenen* Gruppe. Das neutrale Element einer additiven Gruppe wird *Null* genannt und mit 0 bezeichnet. Das inverse Element g' wird dann mit $-g$ bezeichnet. Entsprechend heißt $-g$ das *negative Element* zu g.

Beispiele 21.2.

(1) $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$.

(2) $(\mathbb{R}^n,+)$, $(\mathbb{R}^{n\times m},+)$.

(3) $(\mathbb{Q}^\times,\cdot)$, $(\mathbb{R}^\times,\cdot)$.

(4) Die Menge $\mathrm{GL}(n,\mathbb{R})$ der invertierbaren reellen $(n\times n)$-Matrizen mit der Matrixmultiplikation $\circ$ als Gruppenoperation ist eine Gruppe. Die Einheitsmatrix E_n ist das neutrale Element.

(5) Die Gruppe $(\mathrm{GL}(n,\mathbb{R}),\circ)$ heißt die *allgemeine lineare Gruppe*. Für $n\geq 2$ ist $(\mathrm{GL}(n,\mathbb{R}),\circ)$ nicht-kommutativ.

(6) $(\mathrm{L}_1^{\llcorner}(n,\mathbb{R}),\circ)$, $\mathrm{D}^\times(n,\mathbb{R})$, $\mathrm{Perm}(n,\mathbb{R})$, $(\mathrm{U}_\times^{\urcorner}(n,\mathbb{R}),\circ)$ sind Gruppen. Dabei ist $\circ$ die Matrixmultiplikation und die Einheitsmatrix E_n das neutrale Element.

(7) $(\mathrm{S}_n,\circ)$ mit der Komposition von Abbildungen ist eine Gruppe. Die identische Abbildung $\mathrm{id}_{\{1,\ldots,n\}}$ ist das neutrale Element der *Permuationsgruppe* $(\mathrm{S}_n,\circ)$.

Satz und Definition 21.3. *Seien $(G,\diamond)$ mit $(g,h)\mapsto g\diamond h$ eine Gruppe und $H\subseteq G$ eine nicht-leere Teilmenge. Dann sind* (1), (2), (3) *äquivalent.*

(1) *Für alle $g,h\in H$ gilt $g\diamond h\in H$. Die Menge H ist mit der Abbildung $(g,h)\mapsto g\diamond h$ eine Gruppe.*

(2) $(\forall g,h\in H):\ g\diamond h\in H\,,\ h'\in H\,.$

(3) $(\forall g,h\in H):\ g\diamond h'\in H\,.$

In (2) *und* (3) *ist h' das in G gebildete inverse Element zu $h\in H$. Wenn die äquivalenten Bedingungen* (1), (2) *und* (3) *erfüllt sind, dann gelten* (4) *und* (5).

(4) *Die Einselemente der Gruppen H und G stimmen überein.*

(5) *Für alle $h\in H$ stimmt das in G gebildete inverse Element h^{-1} mit dem in H gebildeten inversen Element überein.*

Wenn eine der äquivalenten Bedingungen (1), (2), (3) *erfüllt ist, dann heißt H eine* Untergruppe *von G. Wir schreiben dann $H\leq G$.*

Beispiele 21.4.

(1) Sei G eine Gruppe. Dann gilt $G\leq G$. Sei e das Einselement von G. Dann gilt $\{e\}\leq G$. Aus $K\leq H\leq G$ folgt $K\leq G$.

(2) $(\mathbb{Z},+)\leq(\mathbb{Q},+)\leq(\mathbb{R},+)\,.$

(3) $(\mathbb{Q}^\times,\cdot)\leq(\mathbb{R}^\times,\cdot)\,.$

(4) $(\mathrm{L}_1^{\llcorner}(n,\mathbb{R}),\circ)\leq(\mathrm{L}_\times^{\llcorner}(n,\mathbb{R}),\circ)\leq(\mathrm{GL}(n,\mathbb{R}),\circ)\,.$

(5) $(\mathrm{D}_1^{\llcorner}(n,\mathbb{R}),\circ)\leq(\mathrm{D}_\times^{\llcorner}(n,\mathbb{R}),\circ)\leq(\mathrm{GL}(n,\mathbb{R}),\circ)\,.$

(6) $(\mathrm{Perm}(n,\mathbb{R}),\circ)\leq(\mathrm{GL}(n,\mathbb{R}),\circ)\,.$

(7) $(\mathrm{U}_1^{\urcorner}(n,\mathbb{R}),\circ)\leq(\mathrm{U}_\times^{\urcorner}(n,\mathbb{R}),\circ)\leq(\mathrm{GL}(n,\mathbb{R}),\circ)\,.$

Satz und Definition 21.5. *Sei* $(G, \diamond, \eta)$ *eine Gruppe und* $S \subseteq G$ *eine beliebige Teilmenge.*

(1) *Der Durchschnitt*

$$\langle S \rangle = \bigcap_{S \subseteq H \leq G} H$$

ist eine Untergruppe von G.

(2) $\langle S \rangle$ *ist die kleinste Untergruppe von* G, *die* S *als Teilmenge enthält.*

(3) $\langle S \rangle$ *besteht aus dem neutralen Element* η *und, wenn* $S \neq \emptyset$ *gilt, aus allen endlichen Produkten der Form*

$$s_1 \diamond \ldots \diamond s_k$$

mit $s_\kappa \in S$ *oder* $s'_\kappa \in S$ *für alle* $\kappa = 1, \ldots, k$ *und alle* $k \in \mathbb{N}$.

(4) *Wenn* $\langle S \rangle = G$ *gilt, dann sagen wir, dass* G *von* S erzeugt *wird.*

(5) *Wenn* $\langle S \rangle = G$ *und* $S \neq \emptyset$ *gelten, dann sagen wir, dass die Elemente von* S *die Gruppe* G erzeugen. *In diesem Falle heißen die Elemente von* S Generatoren *oder* erzeugende Elemente *von* G.

Satz 21.6. *Sei* $n \in \mathbb{R}$. *Die Gruppe* $(\mathrm{GL}(n, \mathbb{R}), \circ, E_n)$ *wird von den Elementarmatrizen*

(1) $E_{i \leftrightarrow j}$ *mit* $i > j$,

(2) $E_{\alpha;i}$ *mit* $\alpha > 1$,

(3) $E_{j;\alpha;i}$ *mit* $\alpha > 0$ *und* $i \neq j$

mit $i, j = 1, \ldots, n$ *erzeugt.*

Definition 21.7. *Seien* G *und* H *Gruppen. Eine Abbildung* $\varphi\colon G \to H$ *heißt ein* Gruppenhomomorphismus, *wenn* $\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$ *für alle* g_1, $g_2 \in G$ *gilt.*

Satz 21.8. *Seien* G *und* H *Gruppen und* $\varphi\colon G \to H$ *ein Gruppenhomomorphismus. Dann gelten:*

(1) $\varphi(e) = e$.

(2) $(\forall g \in G)\colon \varphi(g^{-1}) = (\varphi(g))^{-1}$.

(3) $\ker(\varphi) = \{g \in G \mid \varphi(g) = e\} \leq G$.

(4) φ *ist injektiv genau dann, wenn* $\ker(\varphi) = \{e\}$ *gilt.*

(5) $\mathrm{im}(\varphi) = \{h \in H \mid (\exists g \in G)\colon \varphi(g) = h\} \leq H$.

Beweis. Wir schreiben beide Gruppen multiplikativ. Das neutrale Element bezeichnen wir unterschiedslos mit e. Das inverse Element zu einem Element x aus g oder H bezeichnen wir mit x^{-1}. Anstelle von $(\varphi(g))^{-1}$ schreiben wir $\varphi(g)^{-1}$ für $g \in G$. Diese verkürzten Schreibweisen sind üblich.

Nachweis von (1). Weil φ ein Gruppenhomomorphismus ist, gilt

$$\varphi(e) = \varphi(ee) = \varphi(e)\varphi(e)\,.$$

Wir multiplizieren die beiden Seiten der Gleichung von links mit $\varphi(e)^{-1}$ und nutzen die Assoziativität aus.

$$\begin{aligned} e = \varphi(e)^{-1}\varphi(e) &= \varphi(e)^{-1}\big(\varphi(e)\varphi(e)\big) \\ &= \big(\varphi(e)^{-1}\varphi(e)\big)\,\varphi(e) = e\,\varphi(e) = \varphi(e)\,. \end{aligned}$$

Damit ist (1) bewiesen.

Nachweis von (2). Sei $g \in G$. Weil φ ein Gruppenhomomorphismus ist, gilt

$$\varphi(g^{-1})\varphi(g) = \varphi(g^{-1}g) = \varphi(e) = e\,.$$

Im letzten Schritt haben wir (1) verwendet. Aus der Einzigkeit des inversen Elementes zu $\varphi(g)$ folgt, dass $\varphi(g^{-1}) = \varphi(g)^{-1}$ gilt.

Nachweis von (3). Nach (1) gilt $\varphi(e) = e$. Also ist $\ker(\varphi)$ nicht die leere Menge. Seien g_1, $g_2 \in \ker(\varphi)$. Wegen

$$\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2) = ee = e$$

gilt dann auch $g_1g_2 \in \ker(\varphi)$. Sei $g \in \ker(\varphi)$. Wegen (2) gilt

$$\varphi(g^{-1}) = \varphi(g)^{-1} = e^{-1} = e\,.$$

Also gilt $g^{-1} \in \ker(\varphi)$. Damit ist gezeigt, dass $\ker(\varphi)$ eine Untergruppe von G ist. Wir haben dabei mehrfach verwendet, dass die Abbildung φ ein Gruppenhomomorphismus ist.

Nachweis von (4). Angenommen, es gilt $\ker(\varphi) = \{e\}$. Seien g_1, $g_2 \in G$ mit $\varphi(g_1) = \varphi(g_2)$ gegeben. Weil φ ein Gruppenhomomorphismus ist, folgt

$$\varphi(g_1g_2^{-1}) = \varphi(g_1)\varphi(g_2)^{-1} = \varphi(g_1)\varphi(g_1)^{-1} = e\,.$$

Aus der Voraussetzung ergibt sich $g_1g_2^{-1} = e$. Also gilt $g_1 = g_2$. Nun setzen wir umgekehrt voraus, dass φ injektiv ist. Für $g \in \ker(\varphi)$ gilt

$$\varphi(g) = e = \varphi(e)\,.$$

Dabei haben wir (1) verwendet. Die Injektivität impliziert $g = e$. Damit ist das Kriterium (4) bewiesen.

Nachweis von (5). Wegen $\varphi(e) = e$ ist $\operatorname{im}(\varphi)$ nicht die leere Menge. Zu h_1, $h_2 \in \operatorname{im}(\varphi)$ gibt es g_1, $g_2 \in G$ mit $\varphi(g_k) = h_k$ für $k = 1, 2$. Aus

$$h_1h_2 = \varphi(g_1)\varphi(g_2) = \varphi(g_1g_2)$$

folgt $h_1h_2 \in \operatorname{im}(\varphi)$. Zu $h \in \operatorname{im}(\varphi)$ gibt es $g \in G$ mit $\varphi(g) = h$. Aus

$$h^{-1} = \varphi(g)^{-1} = \varphi(g^{-1})$$

folgt $h^{-1} \in \operatorname{im}(\varphi)$. Damit ist gezeigt, dass $\operatorname{im}(\varphi)$ eine Untergrupe von H ist. $\square$

Definition 21.9. *Sei* $(G, \,.\,, e)$ *eine Gruppe. Sei* $n \in \mathbb{N}$.

(1) *Ein Gruppenhomomorphimus* $\varphi\colon G \to \mathrm{GL}(n, \mathbb{R})$ *heißt eine* Gruppendarstellung *von* G *durch reelle* $(n \times n)$*-Matrizen.*

(2) *Eine Gruppendarstellung* $\varphi\colon G \to \mathrm{GL}(n, \mathbb{R})$ *heißt* treu*, wenn* φ *injektiv ist.*

Beispiel 21.10. Sei $n \in \mathbb{N}$. Dann ist $P\colon \mathrm{S}_n \to \mathrm{GL}(n, \mathbb{R})$ mit

$$P(\pi)e_\nu = e_{\pi(\nu)}$$

für alle $\nu = 1, \ldots, n$ eine treue Gruppendarstellung der Permutationsgruppe S_n. Siehe Satz und Definition 9.8.

22 Ergänzung: Euklidische Bewegungen

Sei $n \in \mathbb{N}$. Der Vektorraum $\mathbb{R}^n$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,\,.\,\rangle$, der euklidischen Norm $\|\,.\,\|$ und der Standardorientierung versehen. Die kanonische Basis $\mathcal{E}_n = \{e_1, \ldots, e_n\}$ ist demach eine positiv orientierte Orthonormalbasis des $\mathbb{R}^n$. Mit $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$ bezeichnen wir den Vektorraum der linearen Abbildungen von $\mathbb{R}^n$ in $\mathbb{R}^n$.

Definition 22.1.

(1) Eine Abbildung $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ heißt eine *Isometrie* oder eine *euklidische Bewegung*, wenn

$$\|\varphi(x) - \varphi(x)\| = \|x - y\|$$

für alle x, $y \in \mathbb{R}^n$ gilt. Mit $\mathrm{E}(n, \mathbb{R})$ bezeichnen wir die Menge der euklidischen Bewegungen des $\mathbb{R}^n$.

(2) Eine Isometrie $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ heißt eine *Starrkörperbewegung* oder eine *spezielle euklidische Bewegung*, wenn sie die Orientierung erhält, das heißt, wenn

$$\det(\varphi(e_1), \ldots, \varphi(e_n)) > 0$$

gilt. Mit $\mathrm{SE}(n, \mathbb{R})$ bezeichnen wir die Menge der speziellen euklidischen Bewegungen des $\mathbb{R}^n$.

Orthogonale Abbildungen und Translationen sowie ihre Kompositionen sind Isometrien. Jedes Paar $(R, a) \in \mathrm{O}(n, \mathbb{R}) \times \mathbb{R}^n$ definiert eine Isometrie $\varphi_{R,a} : \mathbb{R}^n \to \mathbb{R}^n$ mit

$$\varphi_{R,a}(x) = Rx + a\,.$$

Umgekehrt kann jede Isometrie des $\mathbb{R}^n$ auf diese Weise dargestellt werden. Dabei erhält $\varphi_{R,a}$ genau dann die Orientierung, wenn $R \in \mathrm{SO}(n, \mathbb{R})$ gilt. Jede Isometrie, die den Nullpunkt festlässt, ist eine lineare Abbildung.

Satz 22.2. *Es gelten:*

(1) *Sei $\psi \in \mathrm{E}(n, \mathbb{R})$ eine Isometrie mit $\psi(0) = 0$. Dann gilt $\psi \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$.*

(2) $\mathrm{E}(n, \mathbb{R}) = \mathrm{O}(n, \mathbb{R}) \times \mathbb{R}^n$.

(3) $\mathrm{SE}(n, \mathbb{R}) = \mathrm{SO}(n, \mathbb{R}) \times \mathbb{R}^n$.

Beweis. Wir heben ausdrücklich hervor, dass in der Definition der Isometrien keine Linearitätseigenschaften vorausgesetzt werden.

Nachweis von (1). Sei $\psi \in \mathrm{E}(n, \mathbb{R})$ eine Isometrie mit $\psi(0) = 0$. Zuerst zeigen wir, dass ψ die euklidische Norm und das euklidische innere Produkt erhält.

Seien x, $y \in \mathbb{R}^n$ beliebig gegeben. Aus den Eigenschaften von ψ folgt

$$\|\psi(x)\| = \|\psi(x) - \psi(0)\| = \|x - 0\| = \|x\| \,.$$

Nach Definition der euklidischen Norm gilt

$$2\,\langle x, y\rangle = \|x\|^2 + \|y\|^2 - \|x - y\|^2 \,.$$

Nun folgt

$$\begin{aligned} 2\,\langle \psi(x), \psi(y)\rangle &= \|\psi(x)\|^2 + \|\psi(y)\|^2 - \|\psi(x) - \psi(y)\|^2 \\ &= \|x\|^2 + \|y\|^2 - \|x - y\|^2 = 2\,\langle x, y\rangle \,. \end{aligned}$$

Weil ψ die euklidische Norm und das euklidische innere Produkt erhält, bilden die Vektoren $\psi(e_1), \ldots, \psi(e_n)$ eine Orthonormalbasis des $\mathbb{R}^n$.

Sei $x \in \mathbb{R}^n$ beliebig gegeben. Dann gibt es eindeutig bestimmte Koeffizienten $x_1, \ldots, x_n \in \mathbb{R}$ und $x'_1, \ldots, x'_n \in \mathbb{R}$ mit

$$x = \sum_{\nu=1}^{n} x_\nu e_\nu \,, \quad \psi(x) = \sum_{\nu=1}^{n} x'_\nu \psi(e_\nu) \,,$$

Weil ψ das euklidische innere Produkt erhält, folgt

$$x'_\nu = \langle \psi(x), \psi(e_\nu)\rangle = \langle x, e_\nu\rangle = x_\nu$$

für alle $\nu = 1, \ldots n$. Siehe Satz 5.1. Also gilt

$$\psi\left(\sum_{\nu=1}^{n} x_\nu e_\nu\right) = \sum_{\nu=1}^{n} x_\nu \psi(e_\nu) \,.$$

Dies zeigt, dass die Abbildung ψ linear ist.

Nachweis von (2) und (3). Sei $\varphi \in \mathrm{E}(n, \mathbb{R})$ eine beliebige Isometrie. Nach (1) definiert $\psi(x) = \varphi(x) - \varphi(0)$ eine lineare Abbildung $\psi \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$. Die Matrixdarstellung von ψ bezüglich $\mathcal{E}_n$ ist eine orthogonale Matrix $R \in \mathrm{O}(n, \mathbb{R})$. Dabei gilt $R \in \mathrm{SO}(n, \mathbb{R})$ genau dann, wenn ψ und damit φ die Orientierung erhält. Mit $a = \varphi(0)$ erhalten wir die gewünschte Darstellung $\varphi(x) = Rx + a$ für alle $x \in \mathbb{R}^n$. Der Beweis ist damit beendet. □

Mit der Komposition $\circ$ von Abbildungen ist $\mathrm{E}(n, \mathbb{R}) = \mathrm{O}(n, \mathbb{R}) \times \mathbb{R}^n$ eine Gruppe. Dabei gilt

$$(R_1, a_1) \circ (R_2, a_2) = (R_1 R_2, R_1 a_2 + a_1) \,.$$

Die Produktmatrix $R_1 R_2$ ist mit R_1 und R_2 ebenfalls eine orthogonale Matrix. Die Gruppe $(\mathrm{E}(n, \mathbb{R}), \circ)$ heißt die *euklidische Gruppe des* $\mathbb{R}^n$.

Im Allgemeinen sind die euklidischen Bewegungen des $\mathbb{R}^n$ keine linearen Abbildungen. Durch Augmentierung lässt sich eine Matrixdarstellung der Gruppe

$(\mathrm{E}(n,\mathbb{R}),\circ)$ erreichen. Einer euklidischen Bewegung $(R,a) \in \mathrm{O}(n,\mathbb{R}) \times \mathbb{R}^n$ wird die Matrix

$$\begin{pmatrix} R & a \\ 0_{1\times n} & 1 \end{pmatrix} \in \mathrm{GL}(n+1,\mathbb{R})$$

zugeordnet. Für alle x, $y \in \mathbb{R}^n$ sind $y = Rx + a$ und

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} Rx + a \\ 1 \end{pmatrix} = \begin{pmatrix} R & a \\ 0_{1\times n} & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}$$

äquivalent. Also kann die euklidische Gruppe $(\mathrm{E}(n,\mathbb{R}),\circ)$ mit einer Untergruppe der allgemeinen linearen Gruppe $(\mathrm{GL}(n+1,\mathbb{R}),\circ)$ identifiziert werden. Dabei wird die spezielle euklidische Bewegung $(E_n, 0_n)$ mit der Einheitsmatrix E_{n+1} identifiziert.

Satz 22.3.

(1) *Für alle* $(R,a) \in \mathrm{E}(n,\mathbb{R})$ *gilt*

$$\det \begin{pmatrix} R & a \\ 0_{1\times n} & 1 \end{pmatrix} = \det(R)\,.$$

(2) *Es gelten die Untergruppenbeziehungen*

$$(\mathrm{SE}(n,\mathbb{R}),\circ) \leq (\mathrm{E}(n,\mathbb{R}),\circ) \leq (\mathrm{GL}(n+1,\mathbb{R}),\circ)\,.$$

Wir definieren Zentralprojektionen im $\mathbb{R}^3$ auf die Ebene $H_{e_3,1}$ mit dem Nullpunkt als Zentrum. Dazu erinnern wir an Definition 4.1. Für $w \in \mathbb{R}^n$ mit $\|w\| = 1$ und $\delta \geq 0$ ist $H_{w,\delta} \subseteq \mathbb{R}^n$ die Hyperebene mit

$$H_{w,\delta} = \{x \in \mathbb{R}^n \mid \langle x, w\rangle = \delta\}\,.$$

Wir heben hervor, dass Zentralprojektionen nicht in allen Raumpunkten erklärt sind. Daher sprechen wir von Zentralprojektionen *im* $\mathbb{R}^3$.

Definition 22.4. *Sei* $p : \mathbb{R}^3 \setminus H_{e_3,0} \to \mathbb{R}^3$ *die Abbildung mit*

$$p(x) = \frac{1}{\langle x, e_3\rangle}\, x\,.$$

Für einen Raumpunkt $x \in \mathbb{R}^3$*, der nicht in der Ebene* $H_{e_3,0}$ *liegt, ist der Bildpunkt* $p(x)$ *der Schnittpunkt der Ebene* $H_{e_3,1}$ *mit der Geraden durch* x *und den Nullpunkt. Die Abbildung* $p : \mathbb{R}^3 \setminus H_{e_3,0} \to \mathbb{R}^3$ *heißt die* Zentralprojektion im $\mathbb{R}^3$ auf die Ebene $H_{e_3,1}$ mit dem Nullpunkt als Zentrum.

Satz 22.5. *Es sei* $(R, a) \in \mathrm{SE}(3, \mathbb{R})$ *eine spezielle euklidische Bewegung. Für einen beliebigen Raumpunkt* $x \in \mathbb{R}^3$ *setzen wir*

$$x' = Rx + a \,, \quad x'' = Rx' + a \,.$$

Weiter sei $p : \mathbb{R}^3 \setminus H_{e_3,0} \to \mathbb{R}^3$ *die Zentralprojektion im* $\mathbb{R}^3$ *auf die Ebene* $H_{e_3,1}$ *mit dem Nullpunkt als Zentrum.*

(1) *Für alle* $x \in \mathbb{R}^3$ *gilt die Orthogonalitätsrelation*

$$\langle Rx', a \times x'' \rangle = 0 \,.$$

(2) (Epipolar constrained.) *Für alle* $x \in \mathbb{R}^3$ *mit* x', $x'' \notin H_{e_3,0}$ *gilt die Orthogonalitätsrelation*

$$\langle Rp(x'), a \times p(x'') \rangle = 0 \,.$$

Eine Kamera erzeugt ebene Bilder von der Oberfläche eines gegebenen räumlichen Objektes. Eine wichtige Klasse von mathematischen Modellen beschreibt diesen Prozess durch eine Komposition aus einer speziellen euklidischen Bewegung und einer Zentralprojektion. Dabei genügt es, Abbildungen der Form

$$c : x \mapsto p(x')$$

zu betrachten. Wie üblich identifizieren wir eine Kamera mit ihrem mathematischen Modell und nennen eine Abbildung c der genannten Form eine *perspektive Kamera* oder eine *projektive Kamera*. Die Rekonstruktion von (R, a) aus Paaren $(p(x'), p(x''))$ erfolgt unter Beachtung der Orthogonalitätsrelation (2) des Satzes 22.5.

23 Ergänzung: Abbildungen

Definition 23.1. *Seien A, B Mengen.*

(1) *Die Menge aller Paare (a, b) mit $a \in A$ und $b \in B$ heißt das* cartesische Produkt *$A \times B$ von A und B. Dabei gilt $(a, b) = (a', b')$ genau dann, wenn $a = a'$ und $b = b'$ gelten. Wenn A oder B die leere Menge ist, dann ist $A \times B$ ebenfalls die leere Menge.*

(2) *Eine Teilmenge $\varphi \subseteq A \times B$ heißt eine* Abbildung *von A* in *B, wenn es zu jedem $a \in A$ genau ein $b \in B$ mit $(a, b) \in \varphi$ gibt. Wir sagen dann auch, dass φ eine Abbildung von A* nach *B ist. Statt* Abbildung *werden die Ausdrücke* Zuordnung *oder* Funktion *verwendet.*

(3) *Es seien A und B nicht-leere Mengen. Wenn $\varphi \subseteq A \times B$ eine Abbildung ist, dann schreiben wir $b = \varphi(a)$ für alle Paare $(a, b) \in \varphi$. Das Element $\varphi(a) \in B$ heißt das* Bild von $a \in A$ unter φ. *Wir sagen dann, dass dem Element $a \in A$ das Element $\varphi(a) \in B$* zugeordnet *wird. Dafür schreiben wir auch $a \mapsto \varphi(a)$.*

(4) *Wenn $\varphi \subseteq A \times B$ eine Abbildung ist, dann schreiben wir $\varphi\colon A \to B$.*

(5) *Sei $\varphi\colon A \to B$ eine Abbildung. Dann heißt die Menge A der* Definitionsbereich *oder die* Definitionsmenge *von φ. Die Menge B heißt der* Zielbereich *oder die* Zielmenge *von φ.*

(6) *Wenn keine Verwechslungen hinsichtlich des Definitions- und des Zielbereiches einer Abbildung $\varphi \subseteq A \times B$ zu befürchten sind, sagen wir, das φ eine Abbildung ist.*

(7) *Sei $\varphi\colon A \to B$ eine Abbildung. Die Menge*

$$\varphi(A) = \{b \in B \mid (\exists a \in A)\colon (a, b) \in \varphi\}$$

heißt das Bild *oder der* Bildbereich *oder die* Bildmenge *von φ. Wenn A oder B die leere Menge ist, dann ist $\varphi(A)$ ebenfalls leer.*

(8) *Sei $\varphi\colon A \to B$ eine Abbildung. Sei $X \subseteq A$. Die Abbildung $\psi\colon X \to B$ mit*

$$\psi \subseteq \varphi \subseteq A \times B$$

heißt die Einschränkung von φ auf X. *Die Abbildung φ ist eine* Fortsetzung von ψ auf A. *Offenbar gilt*

$$\psi(A) \subseteq \varphi(A)\,.$$

Wir schreiben $\varphi \mid X$ für die Einschränkung von φ auf X. Wenn die Mengen X, A und B nicht-leer sind, dann gilt

$$(\varphi \mid A)(x) = \psi(x) = \varphi(x)$$

für alle $x \in X$.

(9) *Sei* $A \neq \emptyset$. *Die Abbildung* $\mathrm{id}_A \colon A \to A$ *mit* $\mathrm{id}_A(a) = a$ *heißt die* identische Abbildung *oder die* Identität *der Menge* A. *Wir setzen* $\mathrm{id}_\emptyset = \emptyset$.

(10) *Eine Abbildung* $\varphi \colon A \to A$ *heißt eine* Selbstabbildung *der Menge* A.

Definition 23.2. *Es seien* A, B, C *Mengen und* $\varphi \colon A \to B$, $\psi \colon B \to C$ *Abbildungen.*

(1) *Die Abbildung* $\psi \circ \varphi \colon A \to C$ *mit*

$$\psi \circ \varphi = \{(a, c) \in A \times C \mid (\exists b \in B) \colon (a, b) \in \varphi, (b, c) \in \psi\}$$

heißt die Komposition von φ und ψ. *Dabei ist auf die Reihenfolge zu achten. Statt* Komposition *werden auch die Ausdrücke* Hintereinanderschaltung *oder* Hintereinanderausführung *verwendet.*

(2) *Wenn die Mengen* A, B, C *nicht-leer sind, dann gilt*

$$(\psi \circ \varphi)(a) = \psi(\varphi(a))$$

für alle $a \in A$. *Dabei wird zuerst in der Menge* B *das Element* $b = \varphi(a)$ *gebildet. Dann wird in der Menge* C *das Element*

$$c = \psi(b) = \psi(\varphi(a))$$

gebildet.

Definition 23.3. *Seien* A, B *Mengen und* $\varphi : A \to B$ *eine Abbildung.*

(1) φ *heißt* injektiv *oder* eineindeutig, *wenn für alle* $x, y \in A$ *aus* $\varphi(x) = \varphi(y)$ *stets* $x = y$ *folgt. Wir sagen dann auch, dass* φ *eine* eins-zu-eins Abbildung von A nach *oder* in B *ist.*

(2) φ *heißt* surjektiv, *wenn es zu jedem* $b \in B$ *ein* $a \in A$ *mit* $\varphi(a) = b$ *gibt. Wir sagen dann, dass* φ *eine* Abbildung von A auf B *ist.*

(3) φ *heißt* bijektiv, *wenn* φ *injektiv und surjektiv ist. Wir sagen dann auch, dass* φ *eine* bijektive Abbildung von A auf B *ist.*

Beispiele 23.4. Es sei $[0, \infty) = \{x \in \mathbb{R} \mid x \geq 0\}$.

(1) $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ ist weder injektiv noch surjektiv.

(2) $f : [0, \infty) \to \mathbb{R}$ mit $f(x) = x^2$ ist injektiv, aber nicht surjektiv.

(3) $f : \mathbb{R} \to [0, \infty)$ mit $f(x) = x^2$ ist surjektiv, aber nicht injektiv.

(4) $f : [0, \infty) \to [0, \infty)$ mit $f(x) = x^2$ ist bijektiv.

Satz 23.5. *Seien A, B Mengen und $\varphi : A \to B$ eine Abbildung.*

(1) *Sei φ bijektiv. Dann gibt es eine eindeutig bestimmte Abbilding $\psi : B \to A$ mit* (i) *und* (ii)*.*

  (i) $(\forall a \in A): \ \psi(\varphi(a)) = a$.

  (ii) $(\forall b \in b): \ \varphi(\psi(b)) = b$.

  Die Abbildung ψ heißt die Umkehrabbildung *von φ und wird mit φ^{-1} bezeichnet. Es gelten* (iii) *und* (iv)*.*

  (iii) *φ^{-1} ist bijektiv.*

  (iv) $(\varphi^{-1})^{-1} = \varphi$.

(2) *φ ist genau dann bijektiv, wenn es eine Abbildung $\psi : B \to A$ mit* (i) *und* (ii) *aus* (1) *gibt. In diesem Fall sind φ und ψ die Umkehrabbildungen voneinander.*

Beispiele 23.6. *Es seien m, $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$ beliebig gegeben. Sei dann $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ die lineare Abbildung mit $\varphi_A(x) = Ax$.*

(1) *φ_A ist genau dann injektiv, wenn $\ker(A) = \{0\}$ gilt.*

(2) *φ_A ist genau dann surjektiv, wenn $\mathrm{rg}(A) = n$ gilt.*

(3) *Im Fall $n = m$ sind äquivalent:*

  (i) *φ_A ist bijektiv.*

  (ii) $\ker(A) = \{0\}$.

  (iii) $\mathrm{rg}(A) = n$.

  (iv) $A \in \mathrm{GL}(n, \mathbb{R})$.

24 Anhang: Klausuren

TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
Carl-Friedrich-Gauß-Fakultät
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen
Am Fallersleber Tore 1
38100 Braunschweig
Dr. Wolfgang Marten

Wintersemester 2014
19.2.2014, Klausurblatt 1/2

Klausur zum Modul
„Lineare Algebra für Studierende
der Informatik“

Aufgabe 1. Gegeben sei die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 8 & 22 & 33 \\ -4 & -16 & -16 \end{pmatrix}.$$

1.1. Berechnen Sie die LU-Zerlegung von A. $\boxed{3.0}$

1.2. Machen Sie die Probe. $\boxed{1.0}$

Aufgabe 2. Gegeben seien L, $U \in \mathbb{R}^{3\times 3}$ und $b \in \mathbb{R}^3$ mit

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -5 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 11 \\ 32 \\ -53 \end{pmatrix}.$$

2.1. Berechnen Sie $A = LU$. $\boxed{1.0}$

2.2. Berechnen Sie die Determinante $\det(A)$. $\boxed{1.0}$

2.3. Lösen Sie $Ax = b$ mit Hilfe der LU-Zerlegung von A. Machen Sie die Probe. Achten Sie auf die Reihenfolge der Komponenten bei der Rückwärtselimination. $\boxed{3.0}$

Aufgabe 3. Gegeben seien $A = (a_1, a_2, a_3, a_4, a_5) \in \mathbb{R}^{4\times 5}$ und $y \in \mathbb{R}^4$ mit

$$A = \begin{pmatrix} 2 & -2 & 2 & 14 & 0 \\ -2 & -2 & -2 & -22 & -1 \\ -1 & 4 & -1 & -1 & 1 \\ -1 & -2 & -1 & -13 & -1 \end{pmatrix}, \quad y = \begin{pmatrix} 10 \\ -30 \\ 11 \\ -21 \end{pmatrix}.$$

3.1. Berechnen Sie die Treppennormalform $T(A, y)$ der erweiterten Matrix $(A, y) \in \mathbb{R}^{4\times 6}$ mit dem Verfahren von Gauß-Jordan. $\boxed{3.0}$

3.2. Bestimmen Sie die Pivotspalten $a_{i_1}, \dots, a_{i_r}$ von A. Die Pivotspalten bilden eine Basis $\mathcal{A} = \{a_{i_1}, \dots, a_{i_r}\}$ von $\mathrm{im}(A) \subseteq \mathbb{R}^4$. $\boxed{1.0}$

3.3. Zeigen Sie, dass $y \in \mathrm{im}(A)$ gilt. Bestimmen Sie mit Hilfe von $T'(A, y)$ den Koordinatenvektor $y_{\mathcal{A}} \in \mathbb{R}^r$ von y bezüglich $\mathcal{A}$. Machen Sie die Probe. $\boxed{2.0}$

3.4. Bestimmen Sie mit Hilfe von $T''(A)$ eine Basis $\mathcal{N}$ von $\ker(A) \subseteq \mathbb{R}^5$. $\boxed{2.0}$

3.5. Bestimmen Sie mit Hilfe von $T''(A, y)$ alle Lösungen von $Ax = y$. $\boxed{1.0}$

Aufgabe 4. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $v_1, v_2, v_3 \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 0 & 4 & 0 \\ -2 & 6 & 0 \\ -2 & 2 & 4 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

4.1. Zeigen Sie, dass die Vektoren v_1, v_2, v_3 eine Basis des $\mathbb{R}^3$ aus Eigenvektoren der Matrix A bilden. Berechnen Sie den Eigenwert λ_i zum Eigenvektor v_i für $i = 1, 2, 3$. $\boxed{7.0}$

4.2. Gegeben sei das reelle Polynom $p(\lambda) = \lambda^3 - 10\lambda^2 + 30\lambda - 24 \in \mathbb{R}[\lambda]$. Berechnen Sie die Matrix

$$p(A) = A^3 - 10A^2 + 30A - 24E_3$$

mit Hilfe der Diagonaldarstellung

$$A = V \circ \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \circ V^{-1}.$$

Dabei ist $V = (v_1, v_2, v_3)$ die Matrix mit den Spalten v_1, v_2, v_3. Berechnen Sie zuerst die Eigenwerte $p(\lambda_i)$ der Matrix $p(A)$ für $i = 1, 2, 3$. Führen Sie die Polynomauswertung mit dem Horner-Schema durch. $\boxed{5.0}$

Hinweis. Die Aufgaben 5 bis 8 finden Sie auf dem Klausurblatt 2/2.

TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
Carl-Friedrich-Gauß-Fakultät
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen
Dr. Wolfgang Marten

Wintersemester 2014
19.2.2014, Klausurblatt 2/2

Klausur zum Modul „Lineare Algebra für Studierende der Informatik“

Aufgabe 5. Gegeben sei $A \in \mathbb{R}^{2\times 2}$ mit

$$A = \begin{pmatrix} -148 & 120 \\ -180 & 146 \end{pmatrix}.$$

5.1. Berechnen Sie die Determinante $\det(A)$, die Adjunkte $\operatorname{adj}(A)$ und die inverse Matrix A^{-1} von A. $\boxed{3.0}$

5.2. Berechnen Sie die Eigenwerte von A. Sei $\lambda_{\max}$ der größte Eigenwert von A. Berechnen Sie eine Basis $\mathcal{B}$ des Eigenraumes $\ker(A - \lambda_{\max}E_2)$. $\boxed{3.0}$

Aufgabe 6. Gegeben sei $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 2 & -4 & 1 \\ 2 & 0 & 2 \\ -1 & 8 & 4 \end{pmatrix}.$$

Berechnen Sie das charakteristische Polynom $\chi_A(t) = \det(A - tE_3)$, die Determinante $\det(A)$, die Adjunkte $\operatorname{adj}(A)$ und die inverse Matrix A^{-1} mit dem Verfahren von Leverrier-Faddeev. $\boxed{8.0}$

Leverrier-Faddeev. Sei $A \in \mathbb{R}^{n\times n}$. Die Koeffizienten $p_1, \dots, p_n \in \mathbb{R}$ von

$$\chi_A(t) = \det(A - tE_n) = (-1)^n(t^n - p_1t^{n-1} - p_2t^{n-2} - \ldots - p_n) \in \mathbb{R}[t]$$

lassen sich für $k = 1, \dots, n$ sukzessive nach dem folgenden Schema berechnen:

$$B_0 = E_n\,, \quad A_k = AB_{k-1}\,, \quad p_k = \frac{\operatorname{tr}(A_k)}{k}\,, \quad B_k = A_k - p_kE_n\,.$$

Dabei ist $E_n \in \mathbb{R}^{n\times n}$ die Einheitsmatrix. Es gelten die Beziehungen

$$\operatorname{tr}(A) = p_1\,, \quad \det(A) = (-1)^{n-1}p_n\,, \quad \operatorname{adj}(A) = (-1)^{n-1}B_{n-1}\,, \quad B_n = 0\,.$$

Wenn A invertierbar ist, dann gilt

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A) = \frac{1}{p_n}\,B_{n-1}\,.$$

Aufgabe 7. Der Vektorraum $\mathbb{R}^3$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,.\,\rangle$ und der Standard-Orientierung versehen. Es sei $\|\,.\,\|$ die euklidische Norm. Gegeben sei die Drehmatrix $A \in \mathrm{SO}(3,\mathbb{R})$ mit

$$A = \frac{1}{13}\begin{pmatrix} -2 & -3-4\sqrt{6} & 6-2\sqrt{6} \\ -3+4\sqrt{6} & -2 & -6-2\sqrt{6} \\ 6+2\sqrt{6} & -6+2\sqrt{6} & 7 \end{pmatrix}.$$

Sei $v \in \mathbb{R}^3 \setminus \{0\}$ der normierte Drehvektor und $\varphi \in (0,\pi)$ der orientierte Drehwinkel von A bezüglich v.

7.1. Berechnen Sie v und $\sin(\varphi)$. Verwenden Sie den schief-symmetrischen Anteil der Drehmatrix A. $\boxed{3.0}$

7.2. Machen Sie die Probe, indem Sie nachrechnen, dass v ein Eigenvektor von A ist. Welcher Eigenwert gehört zu diesem Eigenvektor? $\boxed{1.0}$

7.3. Berechnen Sie $\cos(\varphi)$. Verwenden Sie die Spurformel. Machen die Probe, indem Sie Teil 7.1 verwenden. $\boxed{2.0}$

Aufgabe 8. Der Vektorraum $\mathbb{R}^3$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,.\,\rangle$ versehen. Es sei $\|\,.\,\|$ die euklidische Norm. Gegeben sei die positive Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = 3\begin{pmatrix} 11 & 8 & -8 \\ 8 & 11 & -8 \\ -8 & -8 & 11 \end{pmatrix}.$$

Es ist 81 ein einfacher und 9 ein zweifacher Eigenwert von A.

8.1. Berechnen Sie die Spektralzerlegung

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda$$

von A. Machen Sie die Probe. $\boxed{7.0}$

8.2. Berechnen Sie die positive Quadratwurzel $A^{\frac{1}{2}}$ mit Hilfe der Spektralzerlegung von A. $\boxed{3.0}$

Hinweis: Die Aufgaben 1 bis 4 finden Sie auf dem Klausurblatt 1/2.

Aufgabe 1. Maximal 4 Punkte.

$$
\begin{array}{rrr|l}
1 & 3 & 4 & \\
8 & 22 & 33 & \\
-4 & -16 & -16 & \\
\hline
1 & 3 & 4 & Z_1 \\
0 & -2 & 1 & Z_2 - (+8)Z_1 \\
0 & -4 & 0 & Z_3 - (-4)Z_1 \\
\hline
1 & 3 & 4 & Z_1 \\
0 & -2 & 1 & Z_2 \\
0 & 0 & -2 & Z_3 - (+2)Z_2
\end{array}
\qquad \boxed{3.0}
$$

$$
LU = \begin{pmatrix} 1 & 0 & 0 \\ 8 & 1 & 0 \\ -4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 4 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 4 \\ 8 & 22 & 33 \\ -4 & -16 & -16 \end{pmatrix} = A\,. \qquad \boxed{1.0}
$$

Aufgabe 2. Maximal 5 Punkte.

$$
LU = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -5 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 2 \\ 9 & 2 & 5 \\ -15 & 4 & -8 \end{pmatrix} = A\,. \qquad \boxed{1.0}
$$

$$
\det(A) = \det(L) \cdot \det(U) = 1 \cdot \det(U) = \det(U) = 3 \cdot 2 \cdot 4 = 24\,. \qquad \boxed{1.0}
$$

$$
\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -5 & 2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 11 \\ 32 \\ -53 \end{pmatrix}, \qquad \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 11 \\ -1 \\ 4 \end{pmatrix}.
$$

$$
y_1 = 11\,, \quad y_2 = 32 - 33 = -1\,, \quad y_3 = -53 + 55 + 2 = 4\,. \qquad \boxed{1.0}
$$

$$
x_3 = 1\,, \quad x_2 = \tfrac{1}{2} \cdot (-1 + 1) = 0\,, \quad x_1 = \tfrac{1}{3} \cdot (11 - 0 - 2) = 3\,. \qquad \boxed{1.0}
$$

$$
Ax = \begin{pmatrix} 3 & 0 & 2 \\ 9 & 2 & 5 \\ -15 & 4 & -8 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 9+0+2 \\ 27+0+5 \\ -45+0-8 \end{pmatrix} = \begin{pmatrix} 11 \\ 32 \\ -53 \end{pmatrix} = b\,. \qquad \boxed{1.0}
$$

Aufgabe 3. Maximal 9 Punkte.

Die Treppennormalform $T(A, y)$ von (A, y).

2	−2	2	14	0	10	
−2	−2	−2	−22	−1	−30	
−1	4	−1	−1	1	11	
−1	−2	−1	−13	−1	−21	
1	−1	1	7	0	5	$\frac{1}{2}Z_1$
0	−4	0	−8	−1	−20	$Z_2 + Z_1$
0	3	0	6	1	16	$Z_3 + \frac{1}{2}Z_1$
0	−3	0	−6	−1	−16	$Z_4 + \frac{1}{2}Z_1$
1	−1	1	7	0	5	Z_1
0	−1	0	−2	0	−4	$Z_2 + Z_3$
0	3	0	6	1	16	Z_3
0	0	0	0	0	0	$Z_4 + Z_3$
Treppennormalform $T(A, y)$						
1	0	1	9	0	9	$Z_1 - Z_2$
0	1	0	2	0	4	$-Z_2$
0	0	0	0	1	4	$Z_3 + 3Z_2$
0	0	0	0	0	0	Z_4
Streichen liefert $T'(A, y)$						
1	0	1	9	0	9	Z_1
0	1	0	2	0	4	Z_2
0	0	0	0	1	4	Z_3
Ergänzen liefert $T''(A, y)$						
1	0	1	9	0	9	Z_1
0	1	0	2	0	4	Z_2
0	0	−1	0	0	0	Ergänzung
0	0	0	−1	0	0	Ergänzung
0	0	0	0	1	4	Z_3

$\boxed{3.0}$

Pivotspalten, Rang.

$$\mathcal{A} = \{a_1, a_2, a_5\}\,, \quad r = \mathrm{rg}(A) = 3\,.$$

$\boxed{1.0}$

Lösbarkeit, Koordinatenvektor.

$$\operatorname{rg}(A,y)=\operatorname{rg}(A)=3\,,\quad T'(A,y)=(T'(A),y_{\mathcal{A}})\,,\quad y_{\mathcal{A}}=\begin{pmatrix}9\\4\\4\end{pmatrix}\in\mathbb{R}^3\,. \qquad \boxed{1.0}$$

Probe.

$$(a_1,a_2,a_5)\,y_{\mathcal{A}}=\begin{pmatrix}2&-2&0\\-2&-2&-1\\-1&4&1\\-1&-2&-1\end{pmatrix}\begin{pmatrix}9\\4\\4\end{pmatrix}=\begin{pmatrix}10\\-30\\11\\-21\end{pmatrix}=y\,. \qquad \boxed{1.0}$$

Die Matrix $T''(A,y)=(T''(A),\xi)\in\mathbb{R}^{5\times 6}$.

$$T''(A,y)=\begin{pmatrix}1&0&1&9&0&9\\0&1&0&2&0&4\\0&0&-1&0&0&0\\0&0&0&-1&0&0\\0&0&0&0&1&4\end{pmatrix}=(e_1^{(5)},e_2^{(5)},v_1,v_2,e_5^{(5)},\xi)\,,\quad \xi=\begin{pmatrix}9\\4\\0\\0\\4\end{pmatrix}.$$

Basis des Kernes.

$$v_1=\begin{pmatrix}1\\0\\-1\\0\\0\end{pmatrix},\quad v_2=\begin{pmatrix}9\\2\\0\\-1\\0\end{pmatrix},\quad \mathcal{N}=\{v_1,v_2\}\,,\quad \ker(A)=[\{v_1,v_2\}]\,. \qquad \boxed{2.0}$$

Allgemeine Lösung von $Ax=y$.

$$x=\xi+\alpha_1v_1+\alpha_2v_2=\begin{pmatrix}9\\4\\0\\0\\4\end{pmatrix}+\alpha_1\begin{pmatrix}1\\0\\-1\\0\\0\end{pmatrix}+\alpha_2\begin{pmatrix}9\\2\\0\\-1\\0\end{pmatrix},\quad \alpha_1,\,\alpha_2\in\mathbb{R}\,. \qquad \boxed{1.0}$$

Aufgabe 4. Maximal 12 Punkte.

$$\begin{pmatrix} 0 & 4 & 0 \\ -2 & 6 & 0 \\ -2 & 2 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 4 \\ 2 & 4 & 4 \\ 2 & 8 & 4 \end{pmatrix} . \qquad \boxed{2.0}$$

$$\lambda_1 = 2\,, \quad \lambda_2 = 4\,, \quad \lambda_3 = 4\,. \qquad \boxed{1.0}$$

$$\begin{array}{rrr|rrr|l}
2 & 1 & 1 & 1 & 0 & 0 & \\
1 & 1 & 1 & 0 & 1 & 0 & \\
1 & 2 & 1 & 0 & 0 & 1 & \\
\hline
1 & 1 & 1 & 0 & 1 & 0 & Z_2 \\
0 & -1 & -1 & 1 & -2 & 0 & Z_1 - 2Z_2 \\
0 & 1 & 0 & 0 & -1 & 1 & Z_3 - Z_2 \\
\hline
1 & 1 & 1 & 0 & 1 & 0 & Z_1 \\
0 & -1 & -1 & 1 & -2 & 0 & Z_2 \\
0 & 0 & -1 & 1 & -3 & 1 & Z_3 + Z_2 \\
\hline
1 & 0 & 0 & 1 & -1 & 0 & Z_1 + Z_2 \\
0 & 1 & 1 & -1 & 2 & 0 & -Z_2 \\
0 & 0 & -1 & 1 & -3 & 1 & Z_3 \\
\hline
1 & 0 & 0 & 1 & -1 & 0 & Z_1 \\
0 & 1 & 0 & 0 & -1 & 1 & Z_2 + Z_3 \\
0 & 0 & 1 & -1 & 3 & -1 & -Z_3
\end{array} \qquad \boxed{4.0}$$

$$\begin{array}{r|r|r|r|r}
 & 1 & -10 & 30 & -24 \\
\hline
2 & & 2 & -16 & 28 \\
\hline
 & 1 & -8 & 14 & 4
\end{array}
\qquad
\begin{array}{r|r|r|r|r}
 & 1 & -10 & 30 & -24 \\
\hline
4 & & 4 & -24 & 24 \\
\hline
 & 1 & -6 & 6 & 0
\end{array} \qquad \boxed{2.0}$$

$$p(A) = V \circ \begin{pmatrix} p(\lambda_1) & 0 & 0 \\ 0 & p(\lambda_2) & 0 \\ 0 & 0 & p(\lambda_3) \end{pmatrix} \circ V^{-1} \qquad \boxed{1.0}$$

$$= \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ -1 & 3 & -1 \end{pmatrix} \qquad \boxed{1.0}$$

$$= \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & -4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 8 & -8 & 0 \\ 4 & -4 & 0 \\ 4 & -4 & 0 \end{pmatrix} . \qquad \boxed{1.0}$$

Aufgabe 5. Maximal 6 Punkte.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -148 & 120 \\ -180 & 146 \end{pmatrix}.$$

$$\begin{aligned} \det(A) = a_{11}a_{22} - a_{21}a_{12} &= (-148) \cdot 146 - (-180) \cdot 120 \\ &= -21608 + 21600 = -8\,. \end{aligned} \qquad \boxed{1.0}$$

$$\operatorname{adj}(A) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \begin{pmatrix} 146 & -120 \\ 180 & -148 \end{pmatrix}. \qquad \boxed{1.0}$$

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A) = -\frac{1}{8} \begin{pmatrix} 146 & -120 \\ 180 & -148 \end{pmatrix}. \qquad \boxed{1.0}$$

$$\begin{aligned} \chi_A(t) &= (-148 - t)(146 - t) + 21600 \\ &= -21608 + 2t + t^2 + 21600 \\ &= t^2 + 2t - 8 \qquad \boxed{1.0} \\ &= (t+4)(t-2)\,. \end{aligned}$$

$$\lambda_{\min} = \lambda_1 = -4\,, \quad \lambda_{\max} = \lambda_2 = 2\,. \qquad \boxed{1.0}$$

$$\begin{array}{rr|l} -150 & 120 & \\ -180 & 144 & \\ \hline 5 & -4 & -\frac{1}{30} Z_1 \\ -30 & 24 & Z_2 - Z_1 \\ \hline 5 & -4 & Z_1 \\ 0 & 0 & Z_2 + 6Z_1 \end{array}$$

$$v = \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \quad \mathcal{B} = \{v\}\,, \quad \ker(A + 2E_2) = [v]\,. \qquad \boxed{1.0}$$

$$Av = \begin{pmatrix} -148 & 120 \\ -180 & 146 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} -592 + 600 \\ -720 + 730 \end{pmatrix} = \begin{pmatrix} 8 \\ 10 \end{pmatrix} = 2 \cdot \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \lambda_{\max} v\,.$$

Aufgabe 6. Maximal 8 Punkte.

$$A_1 = A = \begin{pmatrix} 2 & -4 & 1 \\ 2 & 0 & 2 \\ -1 & 8 & 4 \end{pmatrix}, \quad p_1 = 6\,, \quad B_1 = \begin{pmatrix} -4 & -4 & 1 \\ 2 & -6 & 2 \\ -1 & 8 & -2 \end{pmatrix}. \qquad \boxed{1.0}$$

$$A_2 = \begin{pmatrix} -17 & 24 & -8 \\ -10 & 8 & -2 \\ 16 & -12 & 7 \end{pmatrix}, \quad p_2 = -1\,, \quad B_2 = \begin{pmatrix} -16 & 24 & -8 \\ -10 & 9 & -2 \\ 16 & -12 & 8 \end{pmatrix}. \qquad \boxed{2.0}$$

$$A_3 = \begin{pmatrix} 24 & 0 & 0 \\ 0 & 24 & 0 \\ 0 & 0 & 24 \end{pmatrix}, \quad p_3 = 24\,, \quad B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad \boxed{1.0}$$

$$\chi_A(t) = (-1)^3(t^3 - 6t^2 + t - 24) = -t^3 + 6t^2 - t + 24\,. \qquad \boxed{1.0}$$

$$\det(A) = (-1)^2\, p_3 = 24\,. \qquad \boxed{1.0}$$

$$\operatorname{adj}(A) = (-1)^2\, B_2 = \begin{pmatrix} -16 & 24 & -8 \\ -10 & 9 & -2 \\ 16 & -12 & 8 \end{pmatrix}. \qquad \boxed{1.0}$$

$$A^{-1} = \frac{1}{p_3}\, B_2 = \frac{1}{24} \begin{pmatrix} -16 & 24 & -8 \\ -10 & 9 & -2 \\ 16 & -12 & 8 \end{pmatrix}. \qquad \boxed{1.0}$$

Probe.

$$\begin{aligned} A \circ \operatorname{adj}(A) &= \begin{pmatrix} 2 & -4 & 1 \\ 2 & 0 & 2 \\ -1 & 8 & 4 \end{pmatrix} \begin{pmatrix} -16 & 24 & -8 \\ -10 & 9 & -2 \\ 16 & -12 & 8 \end{pmatrix} \\ &= \begin{pmatrix} 24 & 0 & 0 \\ 0 & 24 & 0 \\ 0 & 0 & 24 \end{pmatrix} = \det(A) \cdot E_3\,. \end{aligned}$$

Aufgabe 7. Maximal 6 Punkte.

$$A = \frac{1}{13}\begin{pmatrix} -2 & -3-4\sqrt{6} & 6-2\sqrt{6} \\ -3+4\sqrt{6} & -2 & -6-2\sqrt{6} \\ 6+2\sqrt{6} & -6+2\sqrt{6} & 7 \end{pmatrix},$$

$$A^t = \frac{1}{13}\begin{pmatrix} -2 & -3+4\sqrt{6} & 6+2\sqrt{6} \\ -3-4\sqrt{6} & -2 & -6+2\sqrt{6} \\ 6-2\sqrt{6} & -6-2\sqrt{6} & 7 \end{pmatrix}.$$

$$\tfrac{1}{2}(A-A^t) = \frac{1}{2\cdot 13}\begin{pmatrix} 0 & -8\sqrt{6} & -4\sqrt{6} \\ 8\sqrt{6} & 0 & -4\sqrt{6} \\ 4\sqrt{6} & 4\sqrt{6} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}.$$

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \frac{4\cdot\sqrt{6}}{2\cdot 13}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \frac{2\cdot\sqrt{6}}{13}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}. \qquad \boxed{1.0}$$

$$\|u\|^2 = \frac{2^2\cdot 6}{13^2}\cdot(1+1+4) = \frac{2^4\cdot 3^2}{13^2}, \quad \|u\| = \frac{2^2\cdot 3}{13} = \frac{12}{13}.$$

$$v = \frac{1}{\|u\|}\,u = \frac{13}{12}\cdot\frac{2\cdot\sqrt{6}}{13}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}. \qquad \boxed{1.0}$$

$$\sin(\varphi) = \|u\| = \frac{12}{13}, \quad \sin^2(\varphi) = \frac{144}{169}. \qquad \boxed{1.0}$$

$$Av = \frac{1}{13\cdot\sqrt{6}}\begin{pmatrix} -2 & -3-4\sqrt{6} & 6-2\sqrt{6} \\ -3+4\sqrt{6} & -2 & -6-2\sqrt{6} \\ 6+2\sqrt{6} & -6+2\sqrt{6} & 7 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$= \frac{1}{13\cdot\sqrt{6}}\begin{pmatrix} 13 \\ -13 \\ 26 \end{pmatrix} = v, \quad \lambda = 1. \qquad \boxed{1.0}$$

$$\mathrm{tr}(A) = 1 + 2\cos(\varphi), \quad \mathrm{tr}(A) = \frac{3}{13}, \quad \cos(\varphi) = \frac{1}{2}\left(\frac{3}{13} - \frac{13}{13}\right) = -\frac{5}{13}. \qquad \boxed{1.0}$$

$$\cos^2(\varphi) + \sin^2(\varphi) = \frac{25}{169} + \frac{144}{169} = \frac{169}{169} = 1. \qquad \boxed{1.0}$$

Zwei Zusätze für die kleinen Übungen. Berechnung der Euler-Winkel.

I. Geometrische Methode mit dem Knotenvektor e_N der Drehmatrix A.

$$A = \frac{1}{13}\begin{pmatrix} -2 & -3-4\sqrt{6} & 6-2\sqrt{6} \\ -3+4\sqrt{6} & -2 & -6-2\sqrt{6} \\ 6+2\sqrt{6} & -6+2\sqrt{6} & 7 \end{pmatrix} \in \mathrm{SO}(3,\mathbb{R}), \quad A^{-1} = A^t .$$

$$A = (a_1, a_2, a_3) = D(e_3,\phi)\circ D(e_1,\theta)\circ D(e_3,\psi).$$

$$e_3 \times a_3 = \frac{1}{13}\begin{pmatrix}0\\0\\1\end{pmatrix} \times \begin{pmatrix} 6-2\sqrt{6} \\ -6-2\sqrt{6} \\ 7 \end{pmatrix} = \frac{1}{13}\begin{pmatrix} 6+2\sqrt{6} \\ 6-2\sqrt{6} \\ 0 \end{pmatrix}.$$

$$e_N = \frac{e_3\times a_3}{\|e_3\times a_3\|} = \frac{1}{2\cdot\sqrt{30}}\begin{pmatrix} 6+2\sqrt{6} \\ 6-2\sqrt{6} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{30}}\begin{pmatrix} 3+\sqrt{6} \\ 3-\sqrt{6} \\ 0 \end{pmatrix}.$$

$$e_N = \begin{pmatrix}\cos(\phi)\\ \sin(\phi)\\ 0\end{pmatrix}, \quad \cos(\phi) = \frac{3+\sqrt{6}}{\sqrt{30}}, \quad \sin(\phi) = \frac{3-\sqrt{6}}{\sqrt{30}}.$$

$$\cos(\theta) = \langle e_3, a_3\rangle = \frac{7}{13}, \quad \sin(\theta) = \|e_3\times a_3\| = \frac{2}{13}\sqrt{30}.$$

$$\begin{pmatrix}\cos(\psi)\\ -\sin(\psi)\\ 0\end{pmatrix} = A^t e_N = \frac{1}{13\cdot\sqrt{30}}\begin{pmatrix} -2 & -3+4\sqrt{6} & 6+2\sqrt{6} \\ -3-4\sqrt{6} & -2 & -6+2\sqrt{6} \\ 6-2\sqrt{6} & -6-2\sqrt{6} & 7 \end{pmatrix}\begin{pmatrix} 3+\sqrt{6} \\ 3-\sqrt{6} \\ 0 \end{pmatrix}$$

$$= \frac{1}{13\cdot\sqrt{30}}\begin{pmatrix} -39+13\sqrt{6} \\ -39-13\sqrt{6} \\ 0 \end{pmatrix}$$

$$= \frac{1}{\sqrt{30}}\begin{pmatrix} -3+\sqrt{6} \\ -3-\sqrt{6} \\ 0 \end{pmatrix}.$$

$$\cos(\psi) = -\frac{3-\sqrt{6}}{\sqrt{30}}, \quad \sin(\psi) = \frac{3+\sqrt{6}}{\sqrt{30}}.$$

II. Direkte Berechnung aus den Einträgen der Drehmatrix A.

$$A = \frac{1}{13} \begin{pmatrix} -2 & -3-4\sqrt{6} & 6-2\sqrt{6} \\ -3+4\sqrt{6} & -2 & -6-2\sqrt{6} \\ 6+2\sqrt{6} & -6+2\sqrt{6} & 7 \end{pmatrix} \in \mathrm{SO}(3,\mathbb{R}), \quad A^{-1} = A^t .$$

$$A = (a_{ij}) = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi) .$$

$$\cos(\theta) = a_{33} = \frac{7}{13} ,$$

$$\sin(\theta) = \sqrt{1 - a_{33}^2} = \sqrt{1 - \frac{49}{169}} = \sqrt{\frac{169-49}{169}} = \frac{2}{13}\sqrt{30} .$$

$$\cos(\phi) = -\frac{a_{23}}{\sqrt{1-a_{33}^2}} = (-1) \cdot \frac{13}{13} \cdot \frac{-6-2\sqrt{6}}{2\sqrt{30}} = \frac{3+\sqrt{6}}{\sqrt{30}} ,$$

$$\sin(\phi) = \frac{a_{13}}{\sqrt{1-a_{33}^2}} = \frac{13}{13} \cdot \frac{6-2\sqrt{6}}{2\sqrt{30}} = \frac{3-\sqrt{6}}{\sqrt{30}} .$$

$$\cos(\psi) = \frac{a_{32}}{\sqrt{1-a_{33}^2}} = \frac{13}{13} \cdot \frac{-6+2\sqrt{6}}{2\sqrt{30}} = -\frac{3-\sqrt{6}}{\sqrt{30}} ,$$

$$\sin(\psi) = \frac{a_{31}}{\sqrt{1-a_{33}^2}} = \frac{13}{13} \cdot \frac{6+2\sqrt{6}}{2\sqrt{30}} = \frac{3+\sqrt{6}}{\sqrt{30}} .$$

Aufgabe 8. Maximal 10 Punkte.

$$\mathrm{tr}(A) = 3 \cdot (11 + 11 + 11) = 99\,, \quad 81 \cdot 1 + 9 \cdot 2 = 99\,.$$

Eigenraum zum einfachen Eigenwert 81.

$$\begin{aligned} A - 81E_3 &= 3 \left\{ \begin{pmatrix} 11 & 8 & -8 \\ 8 & 11 & -8 \\ -8 & -8 & 11 \end{pmatrix} - \begin{pmatrix} 27 & 0 & 0 \\ 0 & 27 & 0 \\ 0 & 0 & 27 \end{pmatrix} \right\} \\ &= 3 \begin{pmatrix} -16 & 8 & -8 \\ 8 & -16 & -8 \\ -8 & -8 & -16 \end{pmatrix} \\ &= 24 \begin{pmatrix} -2 & 1 & -1 \\ 1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix}. \end{aligned}$$

$$\ker(A - 81E_3) = \ker \begin{pmatrix} -2 & 1 & -1 \\ 1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix}.$$

$$\begin{array}{rrr|l} -2 & 1 & -1 & \\ 1 & -2 & -1 & \\ -1 & -1 & -2 & \\ \hline 1 & -2 & -1 & Z_2 \\ 0 & -3 & -3 & Z_1 + 2Z_2 \\ 0 & -3 & -3 & Z_3 + Z_2 \\ \hline 1 & 0 & 1 & Z_1 - \frac{2}{3}Z_2 \\ 0 & 1 & 1 & -\frac{1}{3}Z_2 \\ 0 & 0 & 0 & Z_3 - Z_2 \end{array}$$ $\boxed{3.0}$

$$v = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \|v\|^2 = 3\,, \quad \ker(A - 81E_3) = [v]\,.$$ $\boxed{1.0}$

Spektralprojektionen.

$$P_{81} = \frac{v \otimes v}{\|v\|^2} = \frac{vv^t}{\|v\|^2} = \frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}. \qquad \boxed{1.0}$$

$$P_9 = E_3 - P_{81} = \frac{1}{3}\left\{\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}\right\}$$

$$= \frac{1}{3}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}. \qquad \boxed{1.0}$$

$$\operatorname{tr}(P_{81}) = 1\,, \quad \operatorname{tr}(P_9) = 2\,.$$

Verlangte Probe.

$$\sum_{\lambda \in \sigma(A)} \lambda P_\lambda = 81 P_{81} + 9 P_9 = 27\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} + 3\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 33 & 24 & -24 \\ 24 & 33 & -24 \\ -24 & -24 & 33 \end{pmatrix} = A\,. \qquad \boxed{1.0}$$

Positive Quadratwurzel.

$$A^{\frac{1}{2}} = \sum_{\lambda \in \sigma(A)} \sqrt{\lambda}\, P_\lambda = 9 P_{81} + 3 P_9 \qquad \boxed{1.0}$$

$$= 3\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \qquad \boxed{1.0}$$

$$= \begin{pmatrix} 5 & 2 & -2 \\ 2 & 5 & -2 \\ -2 & -2 & 5 \end{pmatrix}. \qquad \boxed{1.0}$$

$$\operatorname{tr}(A^{\frac{1}{2}}) = 5 + 5 + 5 = 15\,, \quad \sqrt{81} \cdot 1 + \sqrt{9} \cdot 2 = 9 + 6 = 15\,.$$

Notenschlüssel Winf	
15 - 17	4.0
18	3.7
19	3.3
20 - 21	3.0
22	2.7
23	2.3
24 - 25	2.0
26	1.7
27	1.3
28 - 30	1.0

Notenschlüssel Inf/Mewi/Dipl	
30 - 32	4.0
33 - 35	3.7
36 - 38	3.3
39 - 41	3.0
42 - 44	2.7
45 - 47	2.3
48 - 50	2.0
51 - 53	1.7
54 - 56	1.3
57 - 60	1.0

Braunschweig, den 5.2.2014/17.9.2014 WM

TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
Carl-Friedrich-Gauß-Fakultät
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen
Am Fallersleber Tore 1
38100 Braunschweig
Dr. Wolfgang Marten

Sommersemester 2014
11.9.2014, Klausurblatt 1/2

Klausur zum Modul „Lineare Algebra für Studierende der Informatik“

Aufgabe 1. Seien $a = (a_i)_{i=1,\ldots,n} \in \mathbb{R}^n$ und $b = (b_j)_{j=1,\ldots,n} \in \mathbb{R}^n$ gegeben. Dann sei $A \in \mathbb{R}^{n\times n}$ die Matrix mit

$$Ax = \langle x, b\rangle a$$

für alle $x \in \mathbb{R}^n$. Dabei ist $\langle\,.\,,\,.\,\rangle$ das euklidische innere Produkt von $\mathbb{R}^n$.

(1) Charakterisieren Sie diejenigen Paare von Vektoren a, $b \in \mathbb{R}^n$, für die die Matrix A symmetrisch ist. $\boxed{4.0}$

(2) Charakterisieren Sie diejenigen Paare von Vektoren a, $b \in \mathbb{R}^n$, für die die Matrix A symmetrisch und idempotent ist. $\boxed{4.0}$

(Eine Matrix $M \in \mathbb{R}^{n\times n}$ heißt *idempotent*, wenn $M = M^2$ gilt.)

Aufgabe 2. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $b \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 5 & -2 & 4 \\ -45 & 16 & -32 \\ 20 & -2 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} -8 \\ 64 \\ -14 \end{pmatrix}.$$

(1) Berechnen Sie die LU-Zerlegung von A. Machen Sie die Probe. $\boxed{3.0}$

(2) Berechnen Sie $\det(A)$ mit Hilfe von (1). $\boxed{1.0}$

(3) Lösen Sie $Ax = b$ mit Hilfe von (1). Machen Sie die Probe. $\boxed{3.0}$

Aufgabe 3. Gegeben seien $A = (a_1, a_2, a_3, a_4, a_5) \in \mathbb{R}^{4\times 5}$ und $y \in \mathbb{R}^4$ mit

$$A = \begin{pmatrix} 1 & -2 & -8 & -17 & 2 \\ -2 & -2 & -14 & -20 & -1 \\ -1 & 3 & 13 & 26 & 1 \\ 2 & -2 & -6 & -16 & -1 \end{pmatrix}, \quad y = \begin{pmatrix} -21 \\ -18 \\ 24 \\ -14 \end{pmatrix}.$$

(1) Berechnen Sie die Matrix $T''(A, y) \in \mathbb{R}^{5\times 6}$, die durch Streichen und Ergänzen aus der Treppennormalform der erweiterten Matrix $(A, y) \in \mathbb{R}^{4\times 6}$ hervorgeht. Verwenden Sie das Verfahren von Gauß-Jordan. $\boxed{3.0}$

(2) Bestimmen Sie mit Hilfe von (1) alle Lösungen von $Ax = y$. Machen Sie die Probe. $\boxed{3.0}$

Aufgabe 4. Sei $A \in \mathbb{R}^{n\times n}$ gegeben. Sei $q(\lambda) \in \mathbb{R}[\lambda]$ das reelle Polynom mit

$$q(\lambda) = \det(A - \lambda E_n).$$

Dabei ist $E_n \in \mathbb{R}^{n\times n}$ die $(n \times n)$-Einheitsmatrix.

(1) Charakterisieren Sie die reellen Nullstellen des Polynoms $q(\lambda)$ mit Hilfe von geeigneten Teilvektorräumen des $\mathbb{R}^n$. $\boxed{1.0}$

(2) Charakterisieren Sie die reellen Nullstellen von $q(\lambda)$ mit Hilfe geeigneter Vektoren des $\mathbb{R}^n$. $\boxed{1.0}$

Aufgabe 5. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $v_1, v_2, v_3 \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 1 & -4 & -4 \\ -4 & 1 & 4 \\ 4 & -4 & -7 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

(1) Zeigen Sie, dass die Vektoren v_1, v_2, v_3 Eigenvektoren von A sind. $\boxed{1.0}$

(2) Sei $V = (v_1, v_2, v_3)$ die Matrix mit den Spalten v_1, v_2, v_3. Berechnen Sie die inverse Matrix V^{-1}. $\boxed{3.0}$

(3) Berechnen Sie die Matrix

$$p(A) = A^3 + 5A^2 + 4A^1 - 6A^0$$

mit Hilfe der Diagonaldarstellung von A. Berechnen Sie zuerst die Eigenwerte von $p(A)$. Verwenden Sie dazu das Horner-Schema. $\boxed{3.0}$

TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
Carl-Friedrich-Gauß-Fakultät
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen — Sommersemester 2014
Dr. Wolfgang Marten — 11.9.2014, Klausurblatt 2/2

Klausur zum Modul „Lineare Algebra für Studierende der Informatik“

Aufgabe 6. Gegeben sei $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 5 & -2 & 4 \\ 2 & 1 & -1 \\ 1 & 2 & -2 \end{pmatrix}.$$

Berechnen Sie die in (1), (2), (3) genannten Objekte mit dem Verfahren von Leverrier-Faddeev.

(1) Das charakteristische Polynom $\chi_A(t) = \det(A - tE_3)$ der Matrix A. $\boxed{1.0}$

(2) Die Determinante $\det(A)$ von A. $\boxed{1.0}$

(3) Die Adjunkte $\operatorname{adj}(A)$ von A. $\boxed{5.0}$

Aufgabe 7. Der Vektorraum $\mathbb{R}^3$ sei mit dem euklidischen inneren Produkt $\langle\, .\, ,\, .\, \rangle$ und der Standard-Orientierung versehen. Es sei $\|\,.\,\|$ die euklidische Norm. Gegeben sei die Drehmatrix $A \in \mathrm{SO}(3, \mathbb{R})$ mit

$$A = (a_1, a_2, a_3) = D(v, \varphi) = \frac{1}{35} \begin{pmatrix} 26 & -15 & 18 \\ 15 & -10 & -30 \\ 18 & 30 & -1 \end{pmatrix}.$$

(1) Berechnen Sie den normierten Drehvektor $v \in \mathbb{R}^3$ mit Hilfe des schiefsymmetrischen Anteils von A. Achten Sie auf den Vorfaktor. $\boxed{3.0}$

(2) Berechnen Sie Cosinus und Sinus des orientierten Drehwinkels $\varphi \in (0, \pi)$ von A bezüglich v. Verwenden Sie die Spurformel. Machen Sie die Probe. $\boxed{3.0}$

(3) Berechnen Sie den Knotenvektor $e_N \in \mathbb{R}^3$ von A. $\boxed{1.0}$

(4) Es sei

$$A = D(e_3, \phi) \circ D(e_1, \theta) \circ D(e_3, \psi)$$

die Euler-Zerlegung der Drehmatrix A. Berechnen Sie Cosinus und Sinus des Euler-Winkels $\phi \in (0, 2\pi)$. $\boxed{2.0}$

Aufgabe 8. Der Vektorraum $\mathbb{R}^3$ sei mit dem euklidischen inneren Produkt $\langle\,.\,,\,.\,\rangle$ versehen. Es sei $\|\,.\,\|$ die euklidische Norm. Gegeben sei die positive Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \frac{1}{9}\begin{pmatrix} 41 & 32 & 16 \\ 32 & 41 & 16 \\ 16 & 16 & 17 \end{pmatrix}.$$

Sie dürfen ohne Beweis verwenden, dass 9 ein einfacher und 1 ein zweifacher Eigenwert der Matrix A ist.

(1) Berechnen Sie die Spur $\operatorname{tr}(A)$ nach der Definition. Überprüfen Sie das Ergebnis mit Hilfe der angegebenen Eigenwerte. $\boxed{2.0}$

(2) Berechnen Sie die Determinante $\det(A)$ mit Hilfe der Eigenwerte von A. $\boxed{1.0}$

(3) Berechnen Sie die Spektralzerlegung

$$A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$$

von A. Machen Sie die Probe. $\boxed{7.0}$

(4) Berechnen Sie die positive Quadratwurzel $A^{\frac{1}{2}}$ mit Hilfe von (3). $\boxed{2.0}$

(5) Berechnen Sie die Adjunkte $\operatorname{adj}(A)$ mit Hilfe von (2) und (3). $\boxed{2.0}$

Aufgabe 1. Maximal 8 Punkte.

(1)

Voraussetzung. Sei $A = A^t$.

$$(\forall x, y \in \mathbb{R}^n) \quad : \quad \langle x, b\rangle\langle a, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \langle y, b\rangle\langle x, a\rangle \quad .$$

$$(\forall i, j \in \{1, \ldots, n\}) \quad : \quad b_i a_j = b_j a_i \quad .$$

- Sei $b \neq 0$. Dann gibt es $j_0 \in \{1, \ldots, n\}$ mit $b_{j_0} \neq 0$. Damit folgt

  $$(\forall i \in \{1, \ldots, n\}) \quad : \quad a_i = \frac{a_{j_0}}{b_{j_0}} b_i \quad .$$

  Also sind a und b linear abhängig. $\boxed{1.0}$
- Sei $b = 0$. Dann sind a und b linear abhängig. $\boxed{1.0}$

- Wenn A symmetrisch ist, dann sind a und b linear abhängig.

Voraussetzung. Seien a und b linear abhängig.

- Sei $b \neq 0$. Dann gibt es $\lambda \in \mathbb{R}$ mit $a = \lambda b$. Damit folgt

  $$(\forall x, y \in \mathbb{R}^n) \quad : \quad \langle Ax, y\rangle = \lambda\langle x, a\rangle\langle a, y\rangle = \lambda\langle y, a\rangle\langle x, a\rangle = \langle x, Ay\rangle \quad .$$

  Also ist A symmetrisch.
- Sei $b = 0$. Dann gilt $A = 0$. Also ist A symmetrisch. $\boxed{1.0}$

- Wenn a und b linear abhängig sind, dann ist A symmetrisch. $\boxed{1.0}$

(2)

Voraussetzung. Sei A symmetrisch und idempotent.

Nach (1) sind a und b linear abhängig.

- Seien $a \neq 0$ und $b \neq 0$. Dann gibt es $\lambda \neq 0$ mit $a = \lambda b$. Es folgt
$$\lambda \langle b, b \rangle b = Ab = A^2 b = \lambda^2 \langle b, b \rangle^2 b \,.$$
Damit erhalten wir
$$\lambda = \frac{1}{\langle b, b \rangle}\,, \quad a = \frac{1}{\langle b, b \rangle}\, b\,, \quad \langle a, b \rangle = 1\,. \qquad \boxed{1.0}$$
- Sei $a = 0$ oder $b = 0$. Dann gilt
$$\langle a, b \rangle = 0\,. \qquad \boxed{1.0}$$

- Wenn die Matrix A symmetrisch und idempotent ist, dann sind a und b linear abhängig mit $\langle a, b \rangle = 0$ oder $\langle a, b \rangle = 1$.

Voraussetzung. Es seien die Vektoren a und b linear abhängig mit $\langle a, b \rangle = 0$ oder $\langle a, b \rangle = 1$.

- Wenn die Vektoren a und b linear abhängig sind und $\langle a, b \rangle = 0$ gilt, dann verschwindet einer der beiden Vektoren. Daher gilt $A = 0$. Die Nullmatrix ist symmetrisch und idempotent. $\boxed{1.0}$
- Es seien a und b linear abhängig mit $\langle a, b \rangle = 1$. Nach (1) ist A symmetrisch. Außerdem gilt
$$(\forall x \in \mathbb{R}^n) \ : \quad A^2 x = \langle x, b \rangle \langle a, b \rangle a = \langle x, b \rangle a = Ax \ \ .$$
Also ist A auch idempotent. $\boxed{1.0}$

- Wenn a und b linear abhängig sind und $\langle a, b \rangle = 0$ oder $\langle a, b \rangle = 1$ gilt, dann ist A symmetrisch und idempotent.

Erinnerung an die Vorlesung. Für das Kronecker-Produkt $A = a \otimes b$ gelten $(a \otimes b)^t = b \otimes a$ und $\mathrm{tr}(a \otimes b) = \langle a, b \rangle$. Das Kronecker-Produkt A ist genau dann symmetrisch und idempotent, wenn A eine Projektionsmatrix ist. In diesem Fall gilt $\mathrm{rg}(A) = \mathrm{tr}(A) \in \{0, 1\}$. Siehe Aufgabe 8, Teil (3).

Aufgabe 2. Maximal 7 Punkte.

$$\begin{array}{rrr|l} 5 & -2 & 4 & \\ -45 & 16 & -32 & \\ 20 & -2 & 10 & \\ \hline 5 & -2 & 4 & Z_1 \\ 0 & -2 & 4 & Z_2-(-9)Z_1 \\ 0 & 6 & -6 & Z_3-(+4)Z_1 \\ \hline 5 & -2 & 4 & Z_1 \\ 0 & -2 & 4 & Z_2 \\ 0 & 0 & 6 & Z_3-(-3)Z_2 \end{array}$$

$\boxed{2.0}$

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ -9 & 1 & 0 \\ 4 & -3 & 1 \end{pmatrix} \begin{pmatrix} 5 & -2 & 4 \\ 0 & -2 & 4 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 5 & -2 & 4 \\ -45 & 16 & -32 \\ 20 & -2 & 10 \end{pmatrix} = A\,.$$

$\boxed{1.0}$

$$\det(A) = \det(L) \cdot \det(U) = 1 \cdot \det(U) = \det(U) = 5 \cdot (-2) \cdot 6 = -60\,.$$

$\boxed{1.0}$

$$\begin{pmatrix} 1 & 0 & 0 \\ -9 & 1 & 0 \\ 4 & -3 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -8 \\ 64 \\ -14 \end{pmatrix}, \quad \begin{pmatrix} 5 & -2 & 4 \\ 0 & -2 & 4 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -8 \\ -8 \\ -6 \end{pmatrix}.$$

$$y_1 = -8\,, \quad y_2 = 64 + 9 \cdot (-8) = -8\,, \quad y_3 = -14 - 4 \cdot (-8) + 3 \cdot (-8) = -6\,.$$

$\boxed{1.0}$

$$x_3 = -1\,, \quad x_2 = (-\tfrac{1}{2}) \cdot (-8+4) = 2\,, \quad x_1 = (\tfrac{1}{5}) \cdot (-8+4+4) = 0\,.$$

$\boxed{1.0}$

$$Ax = \begin{pmatrix} 5 & -2 & 4 \\ -45 & 16 & -32 \\ 20 & -2 & 10 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -8 \\ 64 \\ -14 \end{pmatrix} = b\,.$$

$\boxed{1.0}$

Aufgabe 3. Maximal 6 Punkte.

(1)

$$
\begin{array}{rrrrr|r|l}
1 & -2 & -8 & -17 & 2 & -21 & \\
-2 & -2 & -14 & -20 & -1 & -18 & \\
-1 & 3 & 13 & 26 & 1 & 24 & \\
2 & -2 & -6 & -16 & -1 & -14 & \\
\hline
1 & -2 & -8 & -17 & 2 & -21 & Z_1 \\
0 & -6 & -30 & -54 & 3 & -60 & Z_2 + 2Z_1 \\
0 & 1 & 5 & 9 & 3 & 3 & Z_3 + Z_1 \\
0 & 2 & 10 & 18 & -5 & 28 & Z_4 - 2Z_1 \\
\hline
1 & -2 & -8 & -17 & 2 & -21 & Z_1 \\
0 & 1 & 5 & 9 & 3 & 3 & Z_3 \\
0 & -6 & -30 & -54 & 3 & -60 & Z_2 \\
0 & 2 & 10 & 18 & -5 & 28 & Z_4 \\
\hline
1 & 0 & 2 & 1 & 8 & -15 & Z_1 + 2Z_2 \\
0 & 1 & 5 & 9 & 3 & 3 & Z_2 \\
0 & 0 & 0 & 0 & 21 & -42 & Z_3 + 6Z_2 \\
0 & 0 & 0 & 0 & -11 & 22 & Z_4 - 2Z_2 \\
\hline
1 & 0 & 2 & 1 & 0 & 1 & Z_1 - \frac{8}{21}Z_3 \\
0 & 1 & 5 & 9 & 0 & 9 & Z_2 - \frac{3}{21}Z_3 \\
0 & 0 & 0 & 0 & 1 & -2 & \frac{1}{21}Z_3 \\
0 & 0 & 0 & 0 & 0 & 0 & Z_4 + \frac{11}{21}Z_3 \\
\hline\hline
1 & 0 & 2 & 1 & 0 & 1 & Z_1 \\
0 & 1 & 5 & 9 & 0 & 9 & Z_2 \\
0 & 0 & -1 & 0 & 0 & 0 & \text{Ergänzung} \\
0 & 0 & 0 & -1 & 0 & 0 & \text{Ergänzung} \\
0 & 0 & 0 & 0 & 1 & -2 & Z_3
\end{array}
$$

$$
T''(A, y) = (a_1'', a_2'', a_3'', a_4'', a_5'', \xi) = \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 1 \\ 0 & 1 & 5 & 9 & 0 & 9 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix}. \qquad \boxed{3.0}
$$

(2)

$$x = \xi + \alpha a_3'' + \beta a_4'' , \quad \alpha \in \mathbb{R} , \quad \beta \in \mathbb{R} .$$ 1.0

$$A\xi = \begin{pmatrix} 1 & -2 & -8 & -17 & 2 \\ -2 & -2 & -14 & -20 & -1 \\ -1 & 3 & 13 & 26 & 1 \\ 2 & -2 & -6 & -16 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \\ 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -21 \\ -18 \\ 24 \\ -14 \end{pmatrix} .$$ 1.0

$$Aa_3'' = \begin{pmatrix} 1 & -2 & -8 & -17 & 2 \\ -2 & -2 & -14 & -20 & -1 \\ -1 & 3 & 13 & 26 & 1 \\ 2 & -2 & -6 & -16 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} ,$$

1.0

$$Aa_4'' = \begin{pmatrix} 1 & -2 & -8 & -17 & 2 \\ -2 & -2 & -14 & -20 & -1 \\ -1 & 3 & 13 & 26 & 1 \\ 2 & -2 & -6 & -16 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

Aufgabe 4. Maximal 2 Punkte.

$$(\forall \lambda \in \mathbb{R}) \ : \quad \det(A - \lambda E_n) = 0 \quad \Leftrightarrow \quad \ker(A - \lambda E_n) \neq \{0\}$$ 1.0

$$\Leftrightarrow \quad (\exists v \in \mathbb{R}^n \setminus \{0\}) \ : \quad Av = \lambda v \ .$$ 1.0

Bemerkung. Die Aufgabe 4 ist eine **Erinnerung an die Vorlesung.** Die beiden Aufgaben 4 und 5 gehören zusammen.

Aufgabe 5. Maximal 7 Punkte.

$$\begin{pmatrix} 1 & -4 & -4 \\ -4 & 1 & 4 \\ 4 & -4 & -7 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -6 & -3 & 1 \\ -3 & -3 & -1 \\ -3 & 0 & 1 \end{pmatrix}.$$

$$\lambda_1 = -3\,, \quad \lambda_2 = -3\,, \quad \lambda_3 = 1\,. \qquad \boxed{1.0}$$

$$\begin{array}{rrr|rrr|l}
2 & 1 & 1 & 1 & 0 & 0 & \\
1 & 1 & -1 & 0 & 1 & 0 & \\
1 & 0 & 1 & 0 & 0 & 1 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & Z_3 \\
0 & 1 & -1 & 1 & 0 & -2 & Z_1 - 2Z_3 \\
0 & 1 & -2 & 0 & 1 & -1 & Z_2 - Z_3 \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & Z_1 \\
0 & 1 & -1 & 1 & 0 & -2 & Z_2 \\
0 & 0 & -1 & -1 & 1 & 1 & Z_3 - Z_2 \\
\hline
1 & 0 & 0 & -1 & 1 & 2 & Z_1 + Z_3 \\
0 & 1 & 0 & 2 & -1 & -3 & Z_2 - Z_3 \\
0 & 0 & 1 & 1 & -1 & -1 & -Z_3
\end{array} \qquad \boxed{3.0}$$

$$\begin{array}{r|r|r|r|r}
 & 1 & 5 & 4 & -6 \\
\hline
-3 & & -3 & -6 & 6 \\
\hline
 & 1 & 2 & -2 & 0
\end{array} \qquad\qquad
\begin{array}{r|r|r|r|r}
 & 1 & 5 & 4 & -6 \\
\hline
1 & & 1 & 6 & 10 \\
\hline
 & 1 & 6 & 10 & 4
\end{array}$$

$$p(A) = V \circ \begin{pmatrix} p(\lambda_1) & 0 & 0 \\ 0 & p(\lambda_2) & 0 \\ 0 & 0 & p(\lambda_3) \end{pmatrix} \circ V^{-1}$$

$$= \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} -1 & 1 & 2 \\ 2 & -1 & -3 \\ 1 & -1 & -1 \end{pmatrix} \qquad \boxed{1.0}$$

$$= \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 4 & -4 & -4 \end{pmatrix} \qquad \boxed{1.0}$$

$$= \begin{pmatrix} 4 & -4 & -4 \\ -4 & 4 & 4 \\ 4 & -4 & -4 \end{pmatrix}. \qquad \boxed{1.0}$$

Zusatz für die kleinen Übungen.

$$\chi_A(t) = (-1)(t+3)^2(t-1) = (-1)(t^3+5t^2+3t-9)\,.$$

$$\chi_A(A) = 0\,, \quad A^3+5A^2+3A^1-9A^0 = 0\,.$$

$$\begin{aligned} p(A) &= (A^3+5A^2+3A^1-9A^0) + (A+3A^0) \\ &= A+3A^0 \\ &= \begin{pmatrix} 1 & -4 & -4 \\ -4 & 1 & 4 \\ 4 & -4 & -7 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 4 & -4 & -4 \\ -4 & 4 & 4 \\ 4 & -4 & -4 \end{pmatrix}. \end{aligned}$$

Aufgabe 6. Maximal 7 Punkte.

Berechnungen nach Leverrier-Faddeev.

$$B_0 = E_n\,, \quad A_k = AB_{k-1}\,, \quad p_k = \frac{\mathrm{tr}(A_k)}{k}\,, \quad B_k = A_k - p_k E_n\,.$$

$$A_1 = A = \begin{pmatrix} 5 & -2 & 4 \\ 2 & 1 & -1 \\ 1 & 2 & -2 \end{pmatrix}, \quad p_1 = 4\,, \quad B_1 = \begin{pmatrix} 1 & -2 & 4 \\ 2 & -3 & -1 \\ 1 & 2 & -6 \end{pmatrix}. \qquad \boxed{1.0}$$

$$A_2 = \begin{pmatrix} 5 & 4 & -2 \\ 3 & -9 & 13 \\ 3 & -12 & 14 \end{pmatrix}, \quad p_2 = 5\,, \quad B_2 = \begin{pmatrix} 0 & 4 & -2 \\ 3 & -14 & 13 \\ 3 & -12 & 9 \end{pmatrix}. \qquad \boxed{2.0}$$

$$A_3 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}, \quad p_3 = 6\,, \quad B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad \boxed{1.0}$$

Auswertung nach Leverrier-Faddeev.

$$\chi_A(t) = \det(A - tE_3) = (-1)^3(t^3-4t^2-5t-6) = -t^3+4t^2+5t+6\,. \qquad \boxed{1.0}$$

$$\det(A) = (-1)^2\,p_3 = 6\,. \qquad \boxed{1.0}$$

$$\mathrm{adj}(A) = (-1)^2\,B_2 = \begin{pmatrix} 0 & 4 & -2 \\ 3 & -14 & 13 \\ 3 & -12 & 9 \end{pmatrix}. \qquad \boxed{1.0}$$

Aufgabe 7. Maximal 9 Punkte.

(1)

$$A = (a_1, a_2, a_3) = \frac{1}{35}\begin{pmatrix} 26 & -15 & 18 \\ 15 & -10 & -30 \\ 18 & 30 & -1 \end{pmatrix} \in \mathrm{SO}(3, \mathbb{R}), \quad A^{-1} = A^t .$$

$$\tfrac{1}{2}(A - A^t) = \frac{1}{7}\begin{pmatrix} 0 & -3 & 0 \\ 3 & 0 & -6 \\ 0 & 6 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} . \qquad \boxed{1.0}$$

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 6 \\ 0 \\ 3 \end{pmatrix} = \frac{3}{7}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} , \quad v = \frac{u}{\|u\|} = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} . \qquad \boxed{2.0}$$

(2)

$$\mathrm{tr}(A) = 1 + 2\cos(\varphi) , \quad \mathrm{tr}(A) = \frac{15}{35} = \frac{3}{7} = 1 - \frac{4}{7} , \quad \cos(\varphi) = -\frac{2}{7} . \qquad \boxed{1.0}$$

$$\sin(\varphi) = \|u\| = \frac{3}{7}\sqrt{5} . \qquad \boxed{1.0}$$

$$\cos^2(\varphi) + \sin^2(\varphi) = \frac{4}{49} + \frac{45}{49} = 1 . \qquad \boxed{1.0}$$

(3)

$$e_3 \times a_3 = \frac{1}{35}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 18 \\ -30 \\ -1 \end{pmatrix} = \frac{1}{35}\begin{pmatrix} 30 \\ 18 \\ 0 \end{pmatrix} = \frac{6}{35}\begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix} .$$

$$e_N = \frac{e_3 \times a_3}{\|e_3 \times a_3\|} = \frac{1}{\sqrt{34}}\begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix} . \qquad \boxed{1.0}$$

(4)

$$e_N = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \\ 0 \end{pmatrix} , \quad \cos(\phi) = \frac{5}{\sqrt{34}} , \quad \sin(\phi) = \frac{3}{\sqrt{34}} . \qquad \boxed{2.0}$$

Zusatz für die kleinen Übungen.

$$\cos(\theta) = \langle e_3, a_3\rangle = -\frac{1}{35}\,, \quad \sin(\theta) = \|e_3 \times a_3\| = \frac{6}{35}\sqrt{34}\,.$$

$$\begin{pmatrix} \cos(\psi) \\ -\sin(\psi) \\ 0 \end{pmatrix} = A^t e_N = \frac{1}{35\cdot\sqrt{34}} \begin{pmatrix} 26 & 15 & 18 \\ -15 & -10 & 30 \\ 18 & -30 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix}$$

$$= \frac{1}{35\cdot\sqrt{34}} \begin{pmatrix} 175 \\ -105 \\ 0 \end{pmatrix}$$

$$= \frac{1}{\sqrt{34}} \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}$$

$$\cos(\psi) = \frac{5}{\sqrt{34}}\,, \quad \sin(\psi) = \frac{3}{\sqrt{34}}\,.$$

$$D(e_3,\phi) \circ D(e_1,\theta) \circ D(e_3,\psi)$$

$$= \frac{1}{34\cdot 35} \begin{pmatrix} 5 & -3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & \sqrt{34} \end{pmatrix} \begin{pmatrix} 35 & 0 & 0 \\ 0 & -1 & -6\sqrt{34} \\ 0 & 6\sqrt{34} & -1 \end{pmatrix} \begin{pmatrix} 5 & -3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & \sqrt{34} \end{pmatrix}$$

$$= \frac{1}{34\cdot 35} \begin{pmatrix} 5 & -3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & \sqrt{34} \end{pmatrix} \begin{pmatrix} 175 & -105 & 0 \\ -3 & -5 & -204 \\ 18\sqrt{34} & 30\sqrt{34} & -\sqrt{34} \end{pmatrix}$$

$$= \frac{1}{34\cdot 35} \begin{pmatrix} 884 & -510 & 812 \\ 510 & -340 & -1020 \\ 18\cdot 34 & 30\cdot 34 & -34 \end{pmatrix}$$

$$= \frac{1}{35} \begin{pmatrix} 26 & -15 & 18 \\ 15 & -10 & -30 \\ 18 & 30 & -1 \end{pmatrix}$$

$$= A\,.$$

Aufgabe 8. Maximal 14 Punkte.

$$A = \frac{1}{9}\begin{pmatrix} 41 & 32 & 16 \\ 32 & 41 & 16 \\ 16 & 16 & 17 \end{pmatrix}.$$

$$\lambda_1 = 9\,, \quad \lambda_2 = \lambda_3 = 1\,.$$

(1)

$$\operatorname{tr}(A) = \frac{41+41+17}{9} = \frac{99}{9} = 11\,. \qquad \boxed{1.0}$$

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \lambda_3 = 9+1+1 = 11\,. \qquad \boxed{1.0}$$

(2)

$$\det(A) = \lambda_1\lambda_2\lambda_3 = 9 \cdot 1 \cdot 1 = 9\,. \qquad \boxed{1.0}$$

(3)

$$\begin{aligned} A - 9E_3 &= \frac{1}{9}\begin{pmatrix} 41 & 32 & 16 \\ 32 & 41 & 16 \\ 16 & 16 & 17 \end{pmatrix} - \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} \\ &= \frac{1}{9}\left\{\begin{pmatrix} 41 & 32 & 16 \\ 32 & 41 & 16 \\ 16 & 16 & 17 \end{pmatrix} - \begin{pmatrix} 81 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix}\right\} \\ &= \frac{1}{9}\begin{pmatrix} -40 & 32 & 16 \\ 32 & -40 & 16 \\ 16 & 16 & -64 \end{pmatrix}. \end{aligned}$$

$$\ker(A - 9E_3) = \ker((\tfrac{9}{8})(A - 9E_3)) = \ker\begin{pmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{pmatrix}.$$

$$
\begin{array}{rrr|l}
-5 & 4 & 2 & \\
4 & -5 & 2 & \\
2 & 2 & -8 & \\
\hline
1 & 1 & -4 & \frac{1}{2}Z_3 \\
-5 & 4 & 2 & Z_1 \\
4 & -5 & 2 & Z_2 \\
\hline
1 & 1 & -4 & Z_1 \\
0 & 9 & -18 & Z_2 + 5Z_1 \\
0 & -9 & 18 & Z_3 - 4Z_1 \\
\hline
1 & 1 & -4 & Z_1 \\
0 & 1 & -2 & \frac{1}{9}Z_2 \\
0 & 0 & 0 & Z_3 + Z_2 \\
\hline
1 & 0 & -2 & Z_1 - Z_2 \\
0 & 1 & -2 & Z_2 \\
0 & 0 & 0 & Z_3
\end{array}
\qquad \boxed{3.0}
$$

$$
v = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \|v\|^2 = 9\,, \quad \ker(A - 9E_3) = [v]\,.
$$

$$
P_9 = \frac{v \otimes v}{\|v\|^2} = \frac{vv^t}{\|v\|^2} = \frac{1}{9}\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix}\,. \qquad \boxed{1.0}
$$

$$
P_1 = E_3 - P_{16} = \frac{1}{9}\left\{\begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} - \begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix}\right\} = \frac{1}{9}\begin{pmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{pmatrix}\,. \qquad \boxed{1.0}
$$

$$
\sum_{\lambda \in \sigma(A)} \lambda P_\lambda = 9P_{16} + P_1 = \frac{9}{9}\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{pmatrix} \qquad \boxed{1.0}
$$

$$
= \frac{1}{9}\begin{pmatrix} 41 & 32 & 16 \\ 32 & 41 & 16 \\ 16 & 16 & 17 \end{pmatrix} = A\,. \qquad \boxed{1.0}
$$

(4)

$$A^{\frac{1}{2}} = \sum_{\lambda \in \sigma(A)} \sqrt{\lambda}\, P_\lambda = 3P_9 + P_1$$

$$= \frac{3}{9}\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{pmatrix} \qquad \boxed{1.0}$$

$$= \frac{1}{9}\begin{pmatrix} 17 & 8 & 4 \\ 8 & 17 & 4 \\ 4 & 4 & 11 \end{pmatrix}. \qquad \boxed{1.0}$$

(5)

$$\operatorname{adj}(A) = \det(A)\, A^{-1} = \left(\prod_{\lambda \in \sigma(A)} \lambda\right)\left(\sum_{\lambda \in \sigma(A)} \frac{1}{\lambda} P_\lambda\right) = P_9 + 9P_1$$

$$= \frac{1}{9}\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix} + \frac{9}{9}\begin{pmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{pmatrix} \qquad \boxed{1.0}$$

$$= \frac{1}{9}\begin{pmatrix} 49 & -32 & -16 \\ -32 & 49 & -16 \\ -16 & -16 & 73 \end{pmatrix}. \qquad \boxed{1.0}$$

Notenschlüssel Winf	
15 - 17	4.0
18	3.7
19	3.3
20 - 21	3.0
22	2.7
23	2.3
24 - 25	2.0
26	1.7
27	1.3
28 - 30	1.0

Notenschlüssel Inf/Mewi/Dipl	
30 - 32	4.0
33 - 35	3.7
36 - 38	3.3
39 - 41	3.0
42 - 44	2.7
45 - 47	2.3
48 - 50	2.0
51 - 53	1.7
54 - 56	1.3
57 - 60	1.0

Zusatz für die kleinen Übungen.

$$B_0 = E_n\,,\quad A_k = AB_{k-1}\,,\quad p_k = \frac{\mathrm{tr}(A_k)}{k}\,,\quad B_k = A_k - p_k E_n\,.$$

$$A_1 = A = \frac{1}{9}\begin{pmatrix} 41 & 32 & 16 \\ 32 & 41 & 16 \\ 16 & 16 & 17 \end{pmatrix}\,,\quad p_1 = 11\,,\quad B_1 = \frac{1}{9}\begin{pmatrix} -58 & 32 & 16 \\ 32 & -58 & 16 \\ 16 & 16 & -82 \end{pmatrix}\,.$$

$$A_2 = \frac{1}{9}\begin{pmatrix} -122 & -32 & -16 \\ -32 & -122 & -16 \\ -16 & -16 & -98 \end{pmatrix}\,,\quad p_2 = -19\,,\quad B_2 = \frac{1}{9}\begin{pmatrix} 49 & -32 & -16 \\ -32 & 49 & -16 \\ -16 & -16 & 73 \end{pmatrix}\,.$$

$$A_3 = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}\,,\quad p_3 = 9\,,\quad B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,.$$

$$\chi_A(t) = (-1)^3(t^3 - 11t^2 + 19t - 9) = (-1)(t-1)^2(t-9)\,.$$

	1	-11	19	-9
1		1	-10	9
	1	-10	9	0

$$t^2 - 10t + 9 = 0 \quad \Leftrightarrow \quad t = 5 \pm \sqrt{25-9} = 5 \pm 4\,.$$

$$\det(A) = (-1)^2\,p_3 = 9\,.$$

$$\mathrm{adj}(A) = (-1)^2\,B_2 = \frac{1}{9}\begin{pmatrix} 49 & -32 & -16 \\ -32 & 49 & -16 \\ -16 & -16 & 73 \end{pmatrix}\,.$$

Braunschweig, 30.8.2014/14.9.2014 WM

TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
Carl-Friedrich-Gauß-Fakultät
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen
Am Fallersleber Tore 1
38100 Braunschweig
Dr. Wolfgang Marten

Wintersemester 2014
19.2.2014, Klausurblatt

Klausur zum Modul
„Lineare Algebra für Studierende
der Wirtschaftsinformatik“

Aufgabe 1. Gegeben sei die Matrix $A \in \mathbb{R}^{3\times 3}$ mit

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 8 & 22 & 33 \\ -4 & -16 & -16 \end{pmatrix}.$$

1.1. Berechnen Sie die LU-Zerlegung von A. $\boxed{3.0}$

1.2. Machen Sie die Probe. $\boxed{1.0}$

Aufgabe 2. Gegeben seien $L, U \in \mathbb{R}^{3\times 3}$ und $b \in \mathbb{R}^3$ mit

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -5 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 11 \\ 32 \\ -53 \end{pmatrix}.$$

2.1. Berechnen Sie $A = LU$. $\boxed{1.0}$

2.2. Berechnen Sie die Determinante $\det(A)$. $\boxed{1.0}$

2.3. Lösen Sie $Ax = b$ mit Hilfe der LU-Zerlegung von A. Machen Sie die Probe. Achten Sie auf die Reihenfolge der Komponenten bei der Rückwärtselimination. $\boxed{3.0}$

Aufgabe 3. Gegeben seien $A = (a_1, a_2, a_3, a_4, a_5) \in \mathbb{R}^{4\times 5}$ und $y \in \mathbb{R}^4$ mit

$$A = \begin{pmatrix} 2 & -2 & 2 & 14 & 0 \\ -2 & -2 & -2 & -22 & -1 \\ -1 & 4 & -1 & -1 & 1 \\ -1 & -2 & -1 & -13 & -1 \end{pmatrix}, \quad y = \begin{pmatrix} 10 \\ -30 \\ 11 \\ -21 \end{pmatrix}.$$

3.1. Berechnen Sie die Treppennormalform $T(A, y)$ der erweiterten Matrix $(A, y) \in \mathbb{R}^{4\times 6}$ mit dem Verfahren von Gauß-Jordan. $\boxed{3.0}$

3.2. Bestimmen Sie die Pivotspalten $a_{i_1}, \dots, a_{i_r}$ von A. Die Pivotspalten bilden eine Basis $\mathcal{A} = \{a_{i_1}, \dots, a_{i_r}\}$ von $\operatorname{im}(A) \subseteq \mathbb{R}^4$. $\boxed{1.0}$

3.3. Zeigen Sie, dass $y \in \operatorname{im}(A)$ gilt. Bestimmen Sie mit Hilfe von $T'(A, y)$ den Koordinatenvektor $y_{\mathcal{A}} \in \mathbb{R}^r$ von y bezüglich $\mathcal{A}$. Machen Sie die Probe. $\boxed{2.0}$

3.4. Bestimmen Sie mit Hilfe von $T''(A)$ eine Basis $\mathcal{N}$ von $\ker(A) \subseteq \mathbb{R}^5$. $\boxed{2.0}$

3.5. Bestimmen Sie mit Hilfe von $T''(A, y)$ alle Lösungen von $Ax = y$. $\boxed{1.0}$

Aufgabe 4. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $v_1, v_2, v_3 \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 0 & 4 & 0 \\ -2 & 6 & 0 \\ -2 & 2 & 4 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

4.1. Zeigen Sie, dass die Vektoren v_1, v_2, v_3 eine Basis des $\mathbb{R}^3$ aus Eigenvektoren der Matrix A bilden. Berechnen Sie den Eigenwert λ_i zum Eigenvektor v_i für $i = 1, 2, 3$. $\boxed{7.0}$

4.2. Gegeben sei das reelle Polynom $p(\lambda) = \lambda^3 - 10\lambda^2 + 30\lambda - 24 \in \mathbb{R}[\lambda]$. Berechnen Sie die Matrix

$$p(A) = A^3 - 10A^2 + 30A - 24E_3$$

mit Hilfe der Diagonaldarstellung

$$A = V \circ \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \circ V^{-1}.$$

Dabei ist $V = (v_1, v_2, v_3)$ die Matrix mit den Spalten v_1, v_2, v_3. Berechnen Sie zuerst die Eigenwerte $p(\lambda_i)$ der Matrix $p(A)$ für $i = 1, 2, 3$. Führen Sie die Polynomauswertung mit dem Horner-Schema durch. $\boxed{5.0}$

TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
Carl-Friedrich-Gauß-Fakultät
Institut *Computational Mathematics*
AG Partielle Differentialgleichungen
Am Fallersleber Tore 1
38100 Braunschweig
Dr. Wolfgang Marten

Sommersemester 2014
11.9.2014, Klausurblatt

Klausur zum Modul „Lineare Algebra für Studierende der Wirtschaftsinformatik“

Aufgabe 1. Seien $a = (a_i)_{i=1,\dots,n} \in \mathbb{R}^n$ und $b = (b_j)_{j=1,\dots,n} \in \mathbb{R}^n$ gegeben. Dann sei $A \in \mathbb{R}^{n\times n}$ die Matrix mit

$$Ax = \langle x, b\rangle a$$

für alle $x \in \mathbb{R}^n$. Dabei ist $\langle\,.\,,\,.\,\rangle$ das euklidische innere Produkt von $\mathbb{R}^n$.

(1) Charakterisieren Sie diejenigen Paare von Vektoren a, $b \in \mathbb{R}^n$, für die die Matrix A symmetrisch ist. $\boxed{4.0}$

(2) Charakterisieren Sie diejenigen Paare von Vektoren a, $b \in \mathbb{R}^n$, für die die Matrix A symmetrisch und idempotent ist. $\boxed{4.0}$

(Eine Matrix $M \in \mathbb{R}^{n\times n}$ heißt *idempotent*, wenn $M = M^2$ gilt.)

Aufgabe 2. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $b \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 5 & -2 & 4 \\ -45 & 16 & -32 \\ 20 & -2 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} -8 \\ 64 \\ -14 \end{pmatrix}.$$

(1) Berechnen Sie die LU-Zerlegung von A. Machen Sie die Probe. $\boxed{3.0}$

(2) Berechnen Sie $\det(A)$ mit Hilfe von (1). $\boxed{1.0}$

(3) Lösen Sie $Ax = b$ mit Hilfe von (1). Machen Sie die Probe. $\boxed{3.0}$

Aufgabe 3. Gegeben seien $A = (a_1, a_2, a_3, a_4, a_5) \in \mathbb{R}^{4\times 5}$ und $y \in \mathbb{R}^4$ mit

$$A = \begin{pmatrix} 1 & -2 & -8 & -17 & 2 \\ -2 & -2 & -14 & -20 & -1 \\ -1 & 3 & 13 & 26 & 1 \\ 2 & -2 & -6 & -16 & -1 \end{pmatrix}, \quad y = \begin{pmatrix} -21 \\ -18 \\ 24 \\ -14 \end{pmatrix}.$$

(1) Berechnen Sie die Matrix $T''(A, y) \in \mathbb{R}^{5\times 6}$, die durch Streichen und Ergänzen aus der Treppennormalform der erweiterten Matrix $(A, y) \in \mathbb{R}^{4\times 6}$ hervorgeht. Verwenden Sie das Verfahren von Gauß-Jordan. $\boxed{3.0}$

(2) Bestimmen Sie mit Hilfe von (1) alle Lösungen von $Ax = y$. Machen Sie die Probe. $\boxed{3.0}$

Aufgabe 4. Sei $A \in \mathbb{R}^{n\times n}$ gegeben. Sei $q(\lambda) \in \mathbb{R}[\lambda]$ das reelle Polynom mit

$$q(\lambda) = \det(A - \lambda E_n).$$

Dabei ist $E_n \in \mathbb{R}^{n\times n}$ die $(n \times n)$-Einheitsmatrix.

(1) Charakterisieren Sie die reellen Nullstellen des Polynoms $q(\lambda)$ mit Hilfe von geeigneten Teilvektorräumen des $\mathbb{R}^n$. $\boxed{1.0}$

(2) Charakterisieren Sie die reellen Nullstellen von $q(\lambda)$ mit Hilfe geeigneter Vektoren des $\mathbb{R}^n$. $\boxed{1.0}$

Aufgabe 5. Gegeben seien $A \in \mathbb{R}^{3\times 3}$ und $v_1, v_2, v_3 \in \mathbb{R}^3$ mit

$$A = \begin{pmatrix} 1 & -4 & -4 \\ -4 & 1 & 4 \\ 4 & -4 & -7 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

(1) Zeigen Sie, dass die Vektoren v_1, v_2, v_3 Eigenvektoren von A sind. $\boxed{1.0}$

(2) Sei $V = (v_1, v_2, v_3)$ die Matrix mit den Spalten v_1, v_2, v_3. Berechnen Sie die inverse Matrix V^{-1}. $\boxed{3.0}$

(3) Berechnen Sie die Matrix

$$p(A) = A^3 + 5A^2 + 4A^1 - 6A^0$$

mit Hilfe der Diagonaldarstellung von A. Berechnen Sie zuerst die Eigenwerte von $p(A)$. Verwenden Sie dazu das Horner-Schema. $\boxed{3.0}$

Literatur

[1] P.-A. Absil, R. Mahony, R. Sepulchre. *Optimization Algorithms on Matrix Manifolds.* Princeton University Press. Princeton, Oxford, 2008.

[2] Adi Ben-Israel, Thomas N. E. Greville. *Generalized Inverses. Theory and Applications.* Second Edition. CMS Books in Mathematics. Canadian Mathematical Society. Springer-Verlag. New York, Berlin, Heidelberg. 2003.

[3] Alan F. Beardon. *Algebra and Geometry.* Cambridge University Press, Cambridge, 2006.

[4] Élie Cartan. *The Theory of Spinors.* Dover Publications, Inc., New York, 1981. The work first appeared in French in 1937 as *Leçons sur la théorie des spineurs (2 Volumes)*, which was printed from Élie Cartan's lectures, gathered and arranged by André Mercier.

[5] Roger Carter, Graeme Segal, Ian MacDonald. *Lectures on Lie Groups and Lie Algebras.* London Mathematical Society. Student Texts, 32. Cambridge University Press, Cambridge, 1995.

[6] Paul Moritz Cohn. *Elements of Linear Algebra.* Chapman & Hale, London, Glasgow, Weinheim, New York, Tokyo, Melbourne, Madras, 1994.

[7] Paul Moritz Cohn. *Algebra I, II, III.* Second Edition. John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1993, 1995, 1991.

[8] Morton L. Curtis. *Matrix Groups.* Universitext. Springer, New York, Heidelberg, Berlin, 1979.

[9] Roberto Cipolla, Peter Giblin. *Visual Motion of Curves and Surfaces.* Cambridge University Press. Cambridge, 2000. Digitally Printed Version 2009.

[10] Charles A. Desoer. *Notes for a Second Course on Linear Systems.* Van Nostrand Reinhold Company, New York, Cincinnati, Toronto, London, Melbourne, 1970.

[11] P. K. Draxl. *Skew Fields.* London Mathematical Society. Lecture Note Series, 81. Cambridge University Press, Cambridge, 1983. Digitally Printed Version 2007.

[12] Harry Dym. *Linear Algebra in Action.* Graduate Studies in Mathematics, Volume 78. American Mathematical Society, Providence, Rhode Island, 2007.

[13] Heinz Eltermann. *Grundlagen der praktischen Matrizenrechnung.* Hochschultaschenbücher für den Ingenieur, Band 434*. Bibliographisches Institut, Mannheim, Wien, Zürich, 1969.

[14] D. K. Faddeev, V. N. Faddeeva. *Computational Methods of Linear Algebra.* W. H. Freeman and Company. San Francisc, London, 1963

[15] Gerd Fischer. *Lineare Algebra.* Zehnte Auflage. Vieweg Studium. Grundkurs Mathematik, Band 17. Vieweg, Braunschweig, Wiesbaden, 1995.

[16] Felix R. Gantmacher. *The Theory of Matrices.* Volume 1 and 2. AMS Chelsea Publishing, 1959. Second AMS printing 2000. American Mathematical Society, Providence, Rhode Island, 2000.

[17] Gene H. Golub, Charles F. van Loan. *Matrix Computations.* North Oxford Academic Publishers, London, 1986.

[18] Wolfgang Gröbner. *Matrizenrechnung.* BI Hochschultaschenbuch, Band 103/103a. Bibliographisches Institut, Mannheim, 1966.

[19] Werner Greub. *Linear Algebra.* Fourth Edition. Graduate Texts in Mathematics, Band 23. Springer-Verlag, New York, 1995.

[20] Richard Hartley, Andrew Zisserman. *Multiple View Geometry in Computer Vision.* Second Edition. Eighth Printing, 2010. Cambridge University Press, Cambridge.

[21] Morris W. Hirsch, Stephen Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Pure and Applied Mathematics. A Series of Monographs and Textbooks. Volume 60. Academic Press, New York, San Francisco, London, 1974.

[22] Gerhard Janssen. *Lineare Algebra und analytische Geometrie.* Vorlesungsmitschrift und Aufgaben von WM, Wintersemester 1997/98 und Sommersemester 1998. Technische Universität Braunschweig.

[23] Nathan Jacobson. *Basis Algebra I, II.* Second Edition. W. H. Freeman and Company, New York, 1985, 1989.

[24] Max Koecher. *Lineare Algebra und analytische Geometrie.* Zweite Auflage. Grundwissen Mathematik, Band 2. Springer, Berlin, Heidelberg, New York, Tokyo, 1985.

[25] Gerhard Kowalewski. *Einführung in die Determinantentheorie einschließlich der unendlichen und der Fredholmschen Determinanten.* AMS Chelsea Publishing. American Mathematical Society. Providence, Rhode Island, 1945.

[26] Hans-Joachim Kowalsky. *Lineare Algebra.* 9. Auflage. de Gruyter Lehrbuch. Walter de Gruyter, Berlin, New York, 1979.

[27] Hans-Joachim Kowalsky, Gerhard O. Michler, *Lineare Algebra.* 11. Auflage. de Gruyter Lehrbuch. Walter de Gruyter, Berlin, New York, 1998.

[28] Jack B. Kuipers. *Quaternions and Rotation Series Sequences: A Primer with Apllications to Orbits, Aerospace, and Virtual Reality.* Princeton University Press, Princeton, 2002.

[29] Serge Lang. *Linear Algebra.* Third Edition. Undergraduate Texts in Mathemathics. Springer-Verlag, New York, 1997.

[30] Serge Lang. *Algebra.* Graduate Texts in Mathematics, Volume 211. Third revisted Edition 2002. Corrected Fourth Printing Edition 2005. Springer, New York.

[31] Peter D. Lax. *Linear Algebra and its Applications.* A Wiley-Interscience Publication. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, Weinheim, 1997.

[32] Katsumi Nomizu. *Fundamentals of Linear Algebra.* McGraw-Hill Book Company. New York, London, Sidney, 1966

[33] *Princeton Companion to Mathematics.* Edited by Timothy Gowors, June Barrow-Green, Imre Leader. Princeton University Press, Princeton, 2008.

[34] Simon Scott. *Traces and Determinants of Elliptic Pseudodifferentialoperators.* Oxford Mathematical Monographs. Oxford University Press, Oxford. 2010.

[35] Emanuel Sperner. *Einführung in die Analytische Geometrie und Algebra.* Teil 1, Teil 2. Fünfte Auflage. Studia Mathematica. Mathematische Lehrbücher. Vandenhoek & Ruprecht, Göttingen, 1961, 1963.

[36] Gilbert W. Stewart. *Matrix Algorithms. Volume I: Basic Decompositions.* SIAM, Philadelphia, 1998.

[37] Gilbert W. Stewart. *Matrix Algorithms. Volume II: Eigensystems.* SIAM, Philadelphia, 2001.

[38] Uwe Storch, Hartmut Wiebe. *Lehrbuch der Mathematik. Band II. Lineare Algebra.* BI-Wissenschaftsverlag, Mannheim, Wien, Zürich, 1990.

[39] Norbert Straumann. *Klassische Mechanik.* Lecture Notes in Physics, Vol. 289. Springer-Verlag, Berlin, Heidelberg, New York, 1987.

[40] Gilbert Strang, Kai Borre. *Linear Algebra, Geodesy, and GPS.* Wellesley-Cambridge Press, Wellesley, 1997.

[41] Lloyd N. Trefethen, David Bau III. *Numerical Linear Algebra.* SIAM, Philadelphia, 1997.

[42] Lloyd N. Trefethen, Mark Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators.* Princeton University Press, Princeton, 2005.

Braunschweig, 21. Oktober 2014 WM